



Ground truth to fake geographies: machine vision and learning in visual practices

Abelardo Gil-Fournier¹ · Jussi Parikka^{1,2}

Received: 13 September 2019 / Accepted: 18 August 2020 / Published online: 7 November 2020
© The Author(s) 2020

Abstract

This article investigates the concept of the ground truth as both an epistemic and technical figure of knowledge that is central to discussions of machine vision and media techniques of visibility. While ground truth refers to a set of remote sensing practices, it has a longer history in operational photography, such as aerial reconnaissance. Building on a discussion of this history, this article argues that ground truth has shifted from a reference to the physical, geographical ground to the surface of the images echoing earlier points raised by philosopher Jean-Luc Nancy that there is a ground of the image that is central to the task of analysis beyond representational practices. Furthermore, building on the practices of pattern recognition, composite imaging, and different interpretational techniques, we discuss contemporary practices of machine learning that mobilizes geographical earth observation datasets for experimental purposes, including tests such as “fake geography” as well as artistic practices, to show how ground truth is operationalized in such contexts of AI and visual arts.

Keywords Remote sensing · Machine vision · Machine learning · Visual culture · Operational image

1 Introduction

“Knowing how to discern a groundless image from an image that is nothing but a blow is an entire art in itself”—Jean-Luc Nancy (2005: 25).

Geographical knowledge starts with how we see, or even more accurately, with the production of images through which we see, observe, analyze, and identify. Images are the supportive instrument for understanding territorial formations, and their mediating role is crucial in establishing the seeing that defines geographical entities of knowledge. This can include the most (seemingly) inconspicuous practices, such as coloring maps or populating them with place and site names. It can include the observation of how everyday life is filled with a variety of forms of geographical knowledge embedded in digital platforms for navigation and other purposes. Geographic information systems are the mainstay

of such practices that emerge through the mobilization of data and electronic communication technologies where the physical and the virtual sign entangle (see Pickles 1994). What has been established by decades of critical research is that the relationship between geography and images is heavily overdetermined: the visual and epistemic systems giving a sense of landscape formations are embedded in multiple social, colonial, gendered, and other forms of representational biases (see Rose 1993; Rogoff 2000; Thrift 2008). What’s more, this complex role of images in geographical knowledge has also given rise to various forms of epistemic transfers that are addressed in media theory: maps are understood as media (Siegert 2011) and cities are mediated by maps that themselves are materially situated as part of multiple layers of technologies new and old (Mattern 2017). Building on this work and related to questions of automation, calculation, and AI techniques, we are interested in how analytical and synthetic knowledge about surfaces of the world—landscapes and territories—shifts to knowledge about the surface of images.

This article focuses on the concept of *ground truth* that has both a technical and symbolic meaning in how it negotiates relations between images, material surfaces (geographical, landscape, territorial) and their entangled relations—to echo Michel Foucault’s phrasing—in various institutional

✉ Abelardo Gil-Fournier
abelardo.gilfourniermartinez@famucz

Jussi Parikka
j.parikka@soton.ac.uk

¹ FAMU, Prague, Czechia

² University of Southampton, Southampton, UK

arrangements of power and knowledge. Starting from ground truth as a grounding figure of knowledge in remote sensing, we build an argument about the synthetic landscapes experimented within current contexts of AI, which we will refer to as “fake geographies” following a term already proposed in computer science research (Xu 2018). Such fabricated and speculative landscapes are intriguing experiments in the creative use of machine learning techniques that deal with, for example, geographical datasets, as they also relate to the figure of the ground truth that is not anymore grounded on Earth, and have become a shifting, ungrounded reference point that has epistemic and aesthetic value. As a contribution to issues of machine vision (as per this special issue), these questions deal with what becomes decipherable as an image and as a landscape in systems that function primarily through data as their input. In short, we are interested in analyzing the shift where the notion of ground truth is no longer specific to the surface of the ground as a geological or geographic reference point. Instead, ground truth becomes read through the “ground” of the synthetic AI images themselves and how datasets are mobilized in machine learning techniques for visual ends.

Jean-Luc Nancy’s (2005) *The Ground of the Image* proposes a similar shift that troubles a rigid distinction between the figure and the ground, the ground and its representation. For Nancy, the image already contains a ground. Even if Nancy’s focus is on classical art history and the philosophy of images that stems from Western art, it becomes a useful reference point for considering the image itself as containing the material ground imprinted onto it but also what is being cut into existence by the image. Nancy writes:

The image does not stand before the ground like a net or a screen. We do not sink; rather, the ground rises to us in the image. The double separation of the image, its pulling away and its cutting out, form both a protection against the ground and an opening onto it. In reality, the ground is not distinct as ground except in the image: without the image, there would only be indistinct adherence. More precisely: in the image, the ground is distinguished by being doubled (Nancy 2005: 13).

This doubling is both philosophical and technical, as our article aims to show. This argument relates essentially to a range of contemporary data and computational techniques and shifts as part of the broader framework of how we engage with questions of AI—such as different machine learning techniques—as part of operational images: images that do not primarily represent but operate in scientific, military, and other technical systems and institutions (see Farocki 2004).

This article is structured around three key points: first, to address the scope of the notion of ground truth, we map the shift from the truth of the ground to pattern recognition as a significant transformation that also relates to questions of machine vision and machine learning techniques (even if we are not able to go into technical specifics in this article). Second, we show how, from the recognition of patterns, we can move to the building of datasets as a relevant part of the infrastructures of ground truth and machine vision. Third, we look at examples of synthetic geographies as experiments that help to understand the ensemble of images in which the ground becomes synthesized with meaningful aesthetic and epistemological consequences.

At stake in our discussion is the claim that ground truth is read from a mass of images, instead of comparatively off the ground. This leads to the question of how these concepts and terms relate to the contemporary situation of machine vision and, more specifically, machine learning and the production of synthetic landscapes, as we argue at the end of the article.

2 Evidential paradigm: from truth on ground to pattern recognition

Ground truth, as used in geography and environmental sciences, designates the information provided by direct observation—usually at the level of the literal ground, the surface of the Earth—in relation to maps, models, and remote sensing technologies. It is a concept that operates by recognizing a distinction between different sources of data, which, by comparison, can be brought to verify certain features of geography. Ground truths emerge on location; they are local, specific, and situated so as to be able to offer a grounding for the network of technologies of sense and location.

Ground truth is premised on calibration as a central feature of remote sensing that includes how distances can be negotiated and become standardized against a set of features that are assumed to stay regular. Hence, ground truth is itself constantly situated in a set of dynamic processes, which can be argued to be also about forms of social and economic power as some discourses in critical geography pointed out in the 1990s (Pickles 1994). Already in this phase of earlier research, it was noted that the computerized environments of geographical systems might fundamentally change the nature of ground truth, where the mediations are becoming distanced from the actual ground as a material and lived environment: “The computer promotes a remote, detached view of the world

as seen through the filter of the computer database. Intimate knowledge of the world recedes into the background of ‘ground truth’ as the computer screen becomes the medium through which the geographer interacts with the world” (Veregin 1994: 100–101).

However, besides the discussions about the computerization and mediatization of material landscapes, the notion of ground truth has found a new milieu in practices linked to machine learning, where it has become part of the standard vocabulary. In these contexts, ground truth refers to the data provided as sample output values to the training and testing phases. That is, it is the set of given outcomes—obtained by any means—used to build the model during the so-called learning process. Ground truth does not distinguish between sources of data but refers to a distinction between the outcomes produced by a model and the data provided as expected values to be compared with. Hence, this incorporation of the ground into the machine learning operations of producing models becomes a central part—of not only this specialist practice—of how we understand image operations in AI culture: as an interplay of images, data, and material environments.

These two different contexts of use of ground truth unveil a significant transition in the role played by images as part of the monitoring complex of remote sensing of the Earth. In Earth Observation and other instrument and discursive practices, the surface becomes a site of grounding particular truths (see, e.g., Bishop 2011). As we will show next, these truths highlight how the image complex has evolved and displaced the Earth as an object of knowledge to Earth-data and datasets of images that constitute the primary reference point.¹

Already as a linguistic term, ground truth can be recognized to contain an oxymoronic heterogeneity. While references to truth locate the concept in the seemingly immaterial space of epistemological values, the ground part of the concept alludes to a presumed tangible substrate of firm evidence included in its multiple uses across different philosophical discourses too. As Caren Kaplan observes, “‘Ground truth’ anchors contemporary preconceptions about physical geography to the comforting solid matter of the earth’s crust” (Kaplan 2018: 34). The idea of witnessing and proximity is closely related to this epistemological trope of ground truth, which thus resonates with Kaplan’s note about the implied solidity of truth, like a permanent and stable geological formation.

But while the rhetoric of epistemological positions is here already heavily laden with different layers, it is tempting to further interrogate how the solidity of the earth’s crust can

be related to the practices of ground truth. Our approach comes from visual studies and critical artistic studies of AI culture while also drawing on material that has investigated questions of surfaces in both screen culture and in architectural and environmental approaches to territories. This means to account for the ways material surfaces become not only the object of systems of perception and analysis—like in aerial photography—but how they become its core element, as the ground truth is subsumed in their operative logic: not only analytical observing but synthetic creation of images from images.

In this regard, it is important to recall that while the notion of ground truth seems to deal with a sort of unmediated set of facts that emerged directly from the earth, the idea of a pure observation has been extensively contested in the domain of science and technology studies (STS). Indeed, as is convincingly shown in many studies and contexts, observations are always theory-dependent and part of a more detailed back and forth movement of comparison and synthesis in contexts of the materiality of epistemic practices (Knorr-Cetina and Mulkay 1983; Hacking 1983; Latour and Woolgar 1986). Similarly, models are embedded in remote sensing from the first point of contact, so to speak, which makes it impossible to maintain a binary between a (data) model and data (capture) through sensors. As Edwards (2013: xiii) puts it: “Today, no collection of signals or observations—even from satellites, which can ‘see’ the whole planet—becomes global in time and space without first passing through a series of data models”. To name an example: in environmental monitoring, the ground truth of measuring instruments in meteorological stations cannot be separated from the weather forecast modeling practices they are part of. Similarly, seismic data is collected as ground truth as it feeds the prediction of seismic movements.

In such domains, ground truth becomes an operation of validating and adjusting a model to a set of facts measured on the ground. While this intervention of the context of the research in the practices of collecting observations has been extensively acknowledged as mentioned above, we would like to emphasize how the notion of ground truth involves an epistemic realm linked to what Carlo Ginzburg (2013) named the “evidential paradigm” which emerged towards the end of the nineteenth century. In other words, ground truth has its own epistemic history as a figure of knowledge.

In the evidential paradigm, the observer—presented under the persona of the detective—is, on the one hand, able to identify a layer of clues and patterns on top of the undifferentiated roughness of matter, and on the other, to produce a plausible reconstruction of events taking place. In this domain, the analytical and the synthetic coalesce. While emerging in a different context than that of visual epistemology or remote sensing, let alone the synthetic technologies of “fake landscapes” in AI techniques that we will turn to

¹ See the notion of data ensemble in Hoelzl and Marie (2014) or the one of image ensembles in Mackenzie and Munster (2019).

later, this reference to Ginzburg helps to draw attention to a common trait that characterizes practices linked to ground truth: the idea that ground speaks through clues, signs, and evidence that a careful observation of the ground as a register is able to distinguish. Among these are also forensic practices that are again part of the contemporary landscape of technical analysis of surfaces (Weizman 2017) and the mediated practices of witnessing (Schuppli 2020), which both seem to carry forward the evidential paradigm even if in more (technical) media-specific ways.

Ground truth applications are found in disciplines such as archeology, paleontology, and forensic research. In such knowledge practices emerging from remote sensing, ground truth is evoked in various contexts: settlement evidence is checked against their images taken from the air (St. Joseph 1945), mass graves are unearthed in relation to the indexical appearance of certain species of plants (Cox, Flavel and Hanson 2008), and agricultural sites are correlated to the detection of phytoliths in soil probes under the microscope (Lombardo et al. 2020). Ground truth is actually a broader term for knowledge verification and calibration that circulates in diverse contexts and practices. In other words, ground truth surfaces as an operation where sets of material traces are distinguished as registers of information. If the detective recognizes and operationalizes the material arrangement of a footprint, the dust on a shoe or the ash of a cigarette as clues indicative of a potential event, in a similar way ground truth encapsulates a set of filtered objects to be mobilized as data.

Ground truth relates closely to Schuppli's (2020) operative concept of *material witness*. This epistemological tool acknowledges not only the evidential role of matter as an active register of external events but also the intervention of explicit acts of scrutiny in the reading of imprints. Here, the "truth" of these "grounds" relies on the application of a series of techniques of demarcation, filtering, and observation, that is, a domain of practitioners and a material culture "that enable such matter to bear witness" (Schuppli 2020: 3). The forensic method of reading material culture acknowledges that ground alone tells no message; hence we are dealing with "impure matter" (Schuppli 2012) affected by the acts of looking. Furthermore, in such contexts of ground truth being established by comparison and other methods, questions of analysis pick up another take on the evidential paradigm as it becomes involved in advanced computational techniques including machine learning—and how it stems from pattern recognition.

Not by chance, since the early 2000s, an educational game by NASA—"Where on Earth...?"—has been inviting players to become "geographical detectives" (California Institute of Technology 2019). The game consists of quizzes where users are asked to locate the geographical area shown on a satellite image by using their abilities to extract visual

clues from the image to recognize the place. The archive of quizzes displays islands, mountain ranges, deltas, volcanos, and other geographical landmarks, pictured from above and presented without any revealing textual key. This case would seem anecdotal if it were not for a recent machine learning project that aims to do something similar on Google's platform, the PlaNet neuronal network (Weyand, Kostrikov, and Philbin 2016). The project aims to build a machine learning model with the ability to determine the location of a photograph by looking at its pixels, that is, without accessing any image metadata such as GPS information.

When comparing NASA's educational game to Google's PlaNet, the task's similarity highlights the main difference: the detective player has been replaced by an algorithmic, big-data-driven process. Beyond the characteristic automatization of pattern recognition in machine learning systems (Mackenzie 2017), we want to address an additional noticeable difference between the two examples. What is interesting here is an assumption underlying the context of the PlaNet project: that any image is supposed to contain in itself enough visual clues and patterns for a sufficiently trained AI model to be able to recognize the place on Earth where it was taken. While for the player in NASA's game, the knowledge of geography linked to her pattern recognition skills is enough to complete the task, in the machine learning project the computational model is supposed to be able to identify the place shown in a photograph—once it has been trained with a large enough dataset of all sorts of outdoor images that are labeled with their geolocation. The physical immutability of the ground in geographical knowledge is replaced by a machine-readable statistical correlation to be found among images when comparing one to another. This is what Adrian Mackenzie and Anna Munster name as the "invisuality" of "platform seeing": "Collections of images operate within and help form a field of distributed invisuality in which relations between images count more than any indexicality or iconicity of an image" (2019: 16). That is, the ground for geographical detectives is replaced by an invisible ground of relations amidst the dataset in the context of statistical learning.

3 Photomosaics and stitching ground truths

Although the term "ground truth" was not used widely until the 1960s,² the topographical techniques of "ground truthing"—that is, synchronizing images and maps as part of epistemic procedures of verification and calibration—as well

² A Google Ngram search showing the use of several wordings for ground truth displays the growing popularity of the term since the 1960s. For more on n-grams see Michel et al. (2011).

as their relation to matters of triangulation as in telemetry, were deployed shortly after the invention of photography. Only ten years after Daguerre's invention, French Army officer Aimé Laussedat produced the first aerial surveys with balloons; five years later, French photographer Nadar filed for a patent on the use of overlapping photos in these surveys (Cosgrove and Fox 2010: 24). Besides the early photography context that is interesting as part of the history of photogrammetry, we want to emphasize the centrality of the early twentieth century and the First World War as far as the operationalization of images about landscapes is concerned. As shown in detail by Saint-Amour (2003, 2011, 2014), after the development of airplanes, the military contexts on photography become instrumental in the image-map complex that shifted the ground of ground truthing.

Besides trained personnel with the fine-tuned capacity to read terrains and images, new technologies supported the task of image comparison and synthetic knowledge. On the one hand, interpreters were considered to be a “highly trained interpretive elite,” often compared to detectives (2003: 356) as they had learned to extract as many visual clues as possible from single aerial photographs. Here the link to the evidential paradigm persists clearly. On the other hand, “a complex technological matrix” (2003: 354) was set up to help in the execution of this task. This matrix included technologies such as the stereoscope, used with pairs of aerial images, the hyperstereoscope, an improved version of the latter relying on the constant speed of planes, which used two pictures separated by a known temporal gap (2003: 360–361), and a specific adaptation of the body and the perceptive skills of the interpreters needed to operate these techniques. With the aid of this reconnaissance matrix—the “deadliest weapon in the war” (2003: 357)—armies were able to distinguish features on the images related to elevation, the third dimension of landscapes, such as differentiating trenches from embankments, as well as seeing even what was hidden underneath bridges and forests (2003: 358). Thus, while the aerial image provides a way of transforming landscapes into readable surfaces (Scott 1999), examples prove that when the dimensions of landscape exceeded the representational and encoding capabilities of isolated pictures, a set of operations involving the use of multiple images simultaneously had to be put into practice.

In addition to making elevations visible—in a way, interpreters accessed “a three-dimensional scale model” of them (Saint-Amour 2003: 358)—other technologies helped to produce pictures of areas of a large-scale that would have otherwise been impossible to portrait in a single shot, such as the areas occupied by trenches. In these cases, different images were stitched together to build a large photomosaic. As Saint-Amour has shown (2014), technicians relied on the appearance of several recognizable objects in the images. As if they were ground truthing the images, they identified

these features as reference objects and used them as anchor points when stitching them together. In other words, techniques of ground truthing had two functions in photomosaics: referential objects were usually tracked to compare and link a particular image to the map of the ground, but here they were also employed to connect images to each other. That is, the ground truthing techniques used to keep images linked as maps that are useful for navigating and reading the surface of the earth also operated to keep images linked to each other. This is particularly relevant in this article's scope, as it shows how the same techniques were used in two different operations. The same techniques used to verify the correlation between data (images) and ground were also in operation to keep hold of the domain of images itself, that is, to keep images connected, not only to the surface of the world but to each other as well. It is these operations of comparison, synthesis, synchronization, calibration that define the scope of ground truth as it emerges as a media technique even before contemporary versions of machine vision and machine learning.

The example of the photomosaic shows how techniques involved in the concept of ground truth were also central to enabling the up-scaling of photographic images by stitching them together. Ground truthing becomes—in this case—relational, traveling from the stabilization of the image in relation to the map, onward to the interweaving of images while keeping them as legitimate geographical tools. The epistemic dimension of what is verifiably “there” is managed through the media techniques mentioned above. Interestingly, this is a relational dimension of the concept of ground truth that has been highlighted elsewhere. Writing about the concept of ground truth as used in the domain of contemporary planetary remote sensing, Jennifer Gabrys has observed how “the ground of ground truth is not, however, the final point of resolution in these sensor environments. Instead, it is a reminder of the constant need to draw connections across phenomena. Ground here is connection and concretization” (Gabrys 2016: 71), which points to the similar traits we have put forward through the examples above.

Continuing on, we move to the question of media techniques of ground truth by discussing another influential project of the past decade, leading us to consider how aerial images are integrated into complex data-synthesizing environments that fluctuate between the visual and invisible. Following Mackenzie and Munster (2019), we argue that this relates to how data is being prepared to be platform-ready and that visual data operates in invisible ways. This relates to the processes of synchronization of visual data, made operative for navigational and other purposes.

4 Environments of images: Google Ground truth

Continuing the recycling of existing terms in data-driven platform contexts, Ground Truth was also the name of one of Google's core projects of the 2010s. First publicly described in *The Atlantic* (Madrigal 2012), the strategic relevance of the availability of detailed and accurate GIS systems in the context of the emergence of all-encompassing digital platforms (Gillespie 2010) fuelled development initiatives. These Google projects aimed to extract data from images at a massive scale, such as the reCAPTCHA project (Strauß 2018: 11). Project Ground Truth focused on creating “accurate and comprehensive map data, by conflating multiple inputs, via algorithms and elbow grease” (Lookingbill and Weiss-Malik 2013). Alongside aerial images, creating access to other combinable inputs—not least of which were Street View cars—was instrumental in offering the epistemologically significant synthesis operating at the back of the map services. The Street View images featured three particularly valuable characteristics: they were regularly updated, accurately geolocated, and displayed map-related information such as traffic signs, street names, and brands' logotypes, among others. The Ground Truth Project has been responsible for the developments geared at reading—as in Optical Character Recognition (OCR)—the information printed in the physical world, pictured afterward on Street View images. Notably, these developments involved much more than software engineering, functioning on the coordinated extraction of vast amounts of hours of human cognitive labor—typical of contemporary AI projects (Crawford and Joler 2018; Ganesh 2020; Joler and Pasquinelli 2020). However, the relevance of the Ground Truth Project in relation to this article is not who or what is in charge of recognizing patterns, but instead, where this activity is performed. Significant for our argument about the machine learning version of ground truth is that the surface of images is the key holder of information. The remarkable aspect of this case is precisely the circulation where data from images in datasets is transferred to the images used as geographical maps. The evidence—the clues and signs—are extracted from the surfaces of the images and projected onto the tiled images that make up the map service.

Furthermore, this circulation presents a version of how the stitching of images mentioned earlier persists as a key trait of image analysis and machine vision from aerial images to a multitude of data-points where the ground becomes not a ground but a shifting set of techniques in which the ground is constantly established and calibrated. In this regard, drawing on Nancy and other sources, Ryan Bishop (2011: 276) argues that “[a]erial visual technologies and aesthetics are almost sole grounded, literally, in the

terrestrial”. The inverted is also true when we consider the role of media that establishes the ground as an epistemologically existing composite: the terrestrial is grounded in the aerial (technologies) and, broadly speaking, in the circulation of images. Aerial photography did, as a matter of fact, persist as a key reference point and infrastructural anchor for the development of remote sensing techniques. Early research on remote sensing in the 1960s shows how the need for geolocating the readings of sensors carried by surveying aircraft—such as spectrometers or radiometers—involved the same hardware as used in aerial photography. Before the availability of satellite-based positioning systems such as GPS, aerial images were the means to geolocate readings of in-flight sensors, thus connecting “the recorded sensor signals to the ground truth visible in or derived from aerial photographs taken in the course of the flight test” (Eppler and Merrill 1969: 665). Aerial photographs were shot simultaneously as the sensor measurements were produced, printing, in some cases, the image of the ground next to the image of the sensor in the same plate (Grossman and Marlatt 1966). Then, be it through the bare ocular inspection of an investigator, a computer-aided “photointerpreter” with a light-pen, or the operations of an automated system, all of the sensor data was printed on top of a map or an aerial image of the surveyed zone (Eppler and Merrill 1969) as if it were a layer of geolocated data in a geographical information system. Not by accident, these were works published simultaneous to Ian McHarg's seminal book, *Design with Nature* (1991), which proposed the layer-cake model, acknowledged as a forerunner of GIS (Steiner and Fleming 2019: 173).

The discussion in geography concerning the mediated “remote, detached view of the world” in computational geographical databases (Veregin 1994: 100–101) resonates again in the context of image circulations that precede GIS systems. The use of aerial photographs as a geolocating tool “projects an image of a de-materializing world” (Virilio 1994: 13), where spatio-temporal coordinates are replaced with the circulation of images. While the reference to dematerialization is a characteristic part of the 1980s–2000s discourse concerning digital technologies, the way images are being understood is nowadays approached with a focus on the materiality of the media techniques that are formative of this image-complex. The ability to picture the ground and the sphere and needle of a measuring instrument was already used in the first aerial photograph surveys, where each shot of the ground included the image of an altimeter and clock placed under the camera. Framed initially by the altimeter and the clock, the aerial image acquired the role of a navigational tool itself, while later becoming replaced by some of the platforms already mentioned. The main point is, however, the focus on the shift from ground to images to data, which in the current context of experimental media arts and experimental use of geographical datasets, becomes

included in the construction of “fake” geographies that we turn to next.

5 Ground truth and synthetic (“fake”) geographies

In this article, we have analyzed how the concept of ground truth entails a movement from observations practiced at the ground level to operations at the surface of the image. From images produced on aircraft to Google Street View vehicles, the priority of datasets becomes emphasized as a core feature of ground truthing that is tied closely to environments of images. As such, the primacy of images ties two sets of recent contexts of imaging and ground truth in surprising ways that also reveals something fundamental about such operational images and how they are also mobilized in contemporary experiments that further shift the notion of the ground truth to fictitious, and even extraterrestrial, land surfaces.

In extraterrestrial remote sensing, we are faced with image analysis where the lack of access to the ground means a complete absence of “absolute ground truth” (Smyth et al. 1995: 109). Operations such as the exploration of Venus’ surface through the images taken by the Magellan Mission during the 1990s are peculiarly similar to examples such as Google’s PlaNet. A dataset of human-labeled images of planet surfaces—samples of images of craters and other patterns of landscape filtered by expert observers—is separated from the ensemble of images obtained from the mission and distinguished as ground truth for a statistical learning process aimed at classifying, at a massive scale, the geographic features on the surface of the planet (Smyth et al. 1995). In outer space, the ground of extraterrestrial planets emerges from the techniques embedded in the technological infrastructure of orbiting vehicles. Ground truth is reliant on spacecraft systems, just as earlier cultural techniques were carried by travelers and colonizers on their ships (Siegert 2015).

Complementing the complex and costly procedures of extraterrestrial imaging projects and “comparative planetology” (Likavcan 2019), contemporary experimental media arts projects, including work that elaborates calibration such as Geocinema’s *Framing Territories*,³ deal with AI methods too; they produce a version of synthetic “fake” landscapes. For instance, in the context of experimental computer science, the *Satellite Image Spoofing project* (Xu and Zhao

2018), proposes a technique aimed at creating fake datasets of satellite images, just as deep fakes produce the illusion of portraying non-existing faces. Deep fake landscapes shift both the focus of machine vision and AI systems from the individual face and demonstrate that any image surface—face, landscape, earth, or extraterrestrial—can be treated in similar ways and subject to similar considerations that push questions of ground truth off the ground. In a way, *Asunder*, the art installation by Tega Brain, Julian Oliver, and Bengt Sjöln (2019), works in this way too. The machine-learning driven simulation of imaginary (future) landscapes are examples of how a fictional AI Environmental Manager not only observes but reorganizes specific locations on Earth based on existing environmental data assembled into (at times absurd) projections. While this work is more about the variety of assumptions of rationality built into climate models and projections, it also works with the “machine vision” of fabricating images of terraforming.

Also within techniques designated “fake geography,” the relation between the aerial view and its intrinsic calculability is explored in generative artworks where images of lands are merged with algorithmic textures, such as in *Neural Landscape Network* by Gregory Chatonsky (2016) or *Invisible Cities* by Gene Kogan (2016). With similar techniques, Shi Weili’s *Terra Mars* mobilized artificial neural networks (conditional GAN in this case) and trained it with “with topographical data and satellite imagery of Earth.” This model was then applied to “see Mars differently,” to make it look like Earth, as one visual commentary on imaginaries of terraforming and, as per the artist’s own words, creative use of AI technologies. In a similar vein, the *Terraformed Mars* twitter bot by the physicist Casey Handmer (2018) offers images of “simulated terraformed Mars landscapes every six hours” that are based on datasets such as the Mars Orbital Laser Altimeter (MOLA) dataset.

Often, such works are discussed in terms of the creative uses of algorithmic techniques and AI. Instead, we want to highlight how the image environments themselves are generative, based on data and details extracted and mobilized from comparative techniques. Instead of merely creative AI, we want to refer to Russian filmmaker Lev Kuleshov’s concept of “creative geography” from the 1920s, a concept that already formulated the ability to build “unique spatial realities [...] out of shots taken in different geographical locations or at different times” (Bozak 2011: 97). This principle of Soviet montage recalls the importance of the relational space opened when ensembles of images are taken together, interweaved in technical operations, and made explicit by Harun Farocki through his practice on the soft-montage (Farocki 2009; Pantenburg 2017). Hence, the synthetic nature of images and landscapes revolves not merely around current versions of machine vision and the creative use of datasets in different AI techniques, but the longer legacy of

³ On the question of calibration both as a technical and artistic notion, see Geocinema’s work *Framing Territories* (2019) that focuses on remote sensing and science infrastructures of the Digital Belt and Road, and Hito Steyerl’s *How Not to Be Seen: A Fucking Didactic Educational.MOV File* (2013).

how techniques of ground truths afford a synthetic creation of truths that rise “to us in the image,” to return to Nancy’s phrasing. Indeed, Nancy’s (2005) take on the ground being doubled and framed in the images becomes an essential guideline—although we must add a further note that it is not only a doubling but a radical synthetic multiplication of grounds that takes place in images as they are mobilized in massive quantities of datasets.

6 Conclusion

In this article, we addressed a set of imaging practices related to the production of geographical knowledge, and we have focused on an analysis of the techniques as they relate to a broader domain of the image. The aim has been to address the shift that contemporary AI culture operationalizes from the surfaces of the world—such as landscapes and territories—to environments of images. This relation is more than representational, and it has, over a longer period, been conditioned by a range of media techniques, and this relates to a shift we have tracked through ground truthing, where the concept of ground truth has been shown to leave the surface of the earth, to be read through the operations and decoding—and synthetic combining—of image surfaces.

Furthermore, and focusing on the relevance of aerial photography and photomosaics, we have shown how the notion of ground truth, despite not being mentioned in literature before the 1950s, has an epistemic history as a figure of knowledge which can be traced back to those photographic techniques linked to the first aerial surveys. We have contextualized these with what Carlo Ginzburg exposed as an evidential paradigm (2013), which can also be described using Adrian Mackenzie’s words while quoting Hannah Arendt in relation to artificial intelligence: “The crux of the problem rests on the ‘treatment’ or operations that ‘reduce terrestrial sensibilities and movements’ to symbols” (Mackenzie 2017: 53). Following approaches that discuss the role of images in the contexts of machine learning (Mackenzie and Munster 2019), we have emphasized the importance of the invisual and non-representational domain of relations between images as elements of the data ensembles involved.

In this regard, ground truth has been shown as the set of techniques where these symbols are related to each other as a media operation that, in addition to grounding, current geographical systems are also able to give rise to what we have addressed as fake or synthetic geographies. Existing critical work in geography has articulated similar claims, such as John Pickles who, writing on GIS and “Benjamin’s law of assembling images” affirmed: “In this sense, as well as legitimizing claims to verisimilitude, digital mapping signals the end of mapping as evidence for anything, or at least the emergence of a representational economy whose

illusions—Baudrillard tells us—will be so powerful that it won’t be possible to tell what is real and what is not” (Pickles 2004: 159). However, ground truth is an operation that goes beyond geography and has in AI techniques and machine vision its main domain of application, as also demonstrated in relation to artistic practices that address such ideas of the ground as a speculative, calculated, hypothesized entity. Thus, broadly speaking, the discussion also concerns contemporary AI-based image cultures in widespread terms, when it comes to technologies and the institutions of the verification of data—of ground truths.

Acknowledgements Thank you to Elise Hunchuck for her copyediting and feedback on the draft article and to the special issue editors and reviewers for their feedback. This research has also been supported by Czech Science Foundation funded project 19-26865X. “Operational Images and Visual Culture: Media Archeological Investigations”.

Funding The research has been supported by Czech Science Foundation funded project 19-26865X “Operational Images and Visual Culture: Media Archeological Investigations”.

Code availability Not applicable.

Compliance with ethical standard

Conflicts of interest Not applicable.

Availability of data and material Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bishop R (2011) Project ‘Transparent Earth’ and the Autopsy of Aerial Targeting The Visual Geopolitics of the Underground. *Theory Culture Society* 28:270–286. <https://doi.org/10.1177/0263276411424918>
- Bozak PN (2011) *The Cinematic Footprint: Lights, Camera*. Rutgers University Press, New Brunswick NJ, Natural Resources
- Brain T, Oliver J, and Sjölen B (2019) *Asunder – project website*. <https://asunder.earth/>. Accessed 30 June 2020
- California Institute of Technology (2019) *Where on Earth? Quizzes*. MISR: Jet Propulsion Laboratory. <https://misr.jpl.nasa.gov/quizzes/index.cfm>. Accessed 22 June 2020
- Chatonsky G (2016) *Neural Landscape Network*. Author’s website. <https://chatonsky.net/nln/>. Accessed 22 June 2020

- Crawford K, Joler V (2020) Anatomy of an AI System: The Amazon Echo as an anatomical map of human labor, data and planetary resources. <https://anatomyof.ai/>. Accessed 22 June 2020
- Cosgrove D, Fox WL (2010) *Photography and Flight*. Reaktion Books, London
- Cox M, Flavel A, Hanson I (2008) *The Scientific Investigation of Mass Graves: Towards Protocols and Standard Operating Procedures*. Cambridge University Press, Cambridge
- Edwards PN (2013) *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. The MIT Press, Cambridge
- Eppler WG, Merrill RD (1969) Relating remote sensor signals to ground-truth information. *Proc IEEE* 57:665–675. <https://doi.org/10.1109/PROC.1969.7021>
- Farocki H (2004) Phantom Images Public 29:12–22
- Farocki H (2009) Cross Influence / Soft Montage. In: Antje Ehmann, Kodwo Eshun (Eds.), *Harun Farocki. Against What? Against Whom?* Koenig Books, London, pp. 64–79
- Ganesh MI (2020) Intelligence Work. A is for Another: A Dictionary of AI. <https://aisforanother.net/pages/article16.html>. Accessed 22 June 2020
- Gabrys J (2016) *Program Earth: Environmental Sensing Technology and the Making of a Computational Planet*. University of Minnesota Press, Minneapolis
- Geocinema (2019) Framing Territories. Collective's website. <https://geocinema.network/>. Accessed 30 June 2020
- Gillespie T (2010) The politics of 'platforms'. *New Media and Society* 12:347–364. <https://doi.org/10.1177/1461444809342738>
- Ginzburg C (2013) *Clues, Myths, and the Historical Method*. Johns Hopkins University Press, Baltimore
- Grossman RL, Marlatt WE (1966) A method of showing what a radiometer 'sees' during an aircraft survey. *Proc 4th Symp on Remote Sensing of Environment*. University of Michigan, Ann Arbor MI, pp 571–574
- Hacking I (1983) *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge University Press, Cambridge
- Handmer C (2018) Terraformed Mars. Twitter bot. <https://twitter.com/terraformedmars>. Accessed 22 June 2020
- Hoelzl I, Marie R (2014) Google Street View: navigating the operative image. *Visual Studies* 29:261–271. <https://doi.org/10.1080/1472586X.2014.941559>
- Kaplan C (2018) *Aerial Aftermaths: Wartime from Above*. Duke University Press, Durham
- Kogan G (2016) Invisible Cities. Author's website. <https://opendot.github.io/ml4a-invisible-cities/>. Accessed 22 June 2020
- Knorr-Cetina K, Mulkay M (1983) *Science Observed: Perspectives on the Social Study of Science*. SAGE, London
- Latour B, Woolgar S (1986) *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press, Princeton
- Likavčan L (2019) *Introduction to Comparative Planetology*. Strelka Press, Moscow
- Lookingbill A, Weiss-Malik M (2013) Google I/O 2013 - Project Ground Truth: Accurate Maps Via Algorithms and Elbow Grease. Google Developers Youtube Channel. <https://www.youtube.com/watch?v=FsbLEtS0uls>. Accessed 22 June 2020
- Lombardo U, Iriarte J, Hilbert L, Ruiz-Pérez J, Capriles JM, Veit H (2020) Early Holocene crop cultivation and landscape modification in Amazonia. *Nature* 581:190–193. <https://doi.org/10.1038/s41586-020-2162-7>
- Mackenzie A (2017) *Machine Learners: Archaeology of a Data Practice*. MIT Press, Cambridge
- MacKenzie A, Munster A (2019) Platform Seeing: Image Ensembles and Their Invisibilities. *Theory Culture Society*. <https://doi.org/10.1177/0263276419847508>
- Madrigal AC (2012) How Google Builds Its Maps—and What It Means for the Future of Everything. *The Atlantic*. <https://www.theatlantic.com/technology/archive/2012/09/how-google-builds-its-maps-and-what-it-means-for-the-future-of-everything/261913/>. Accessed 22 June 2020
- Mattern S (2017) *Code and Clay, Data and Dirt: Five Thousand Years of Urban Media*. University of Minnesota Press, Minneapolis
- McHarg IL (1991) *Design with Nature*. John Wiley & Sons, New York
- Michel J-B, Shen YK, Aiden AP et al (2011) Quantitative analysis of culture using millions of digitized books. *Science* 331:176–182. <https://doi.org/10.1126/science.1199644>
- Nancy J-L (2005) *The Ground of the Image*. Fordham University Press, New York
- Pantenburg V (2017) Working images: Harun Farocki and the operational image. In: Eder J, Klonk C (eds) *Image Operations: Visual Media and Political Conflict*. Manchester University Press, Manchester, pp 49–62
- Pasquinelli M, Joler V (2020) The Nooscope Manifested: AI as Instrument of Knowledge Extractivism. <https://nooscope.ai/>. Accessed 22 June 2020
- Pickles J (ed) (1994) *Ground Truth: The Social Implications of Geographic Information Systems*. The Guilford Press, New York
- Pickles J (2004) *A History of Spaces: Cartographic Reason. Mapping and the Geo-Coded World*, Routledge
- Rogoff I (2000) *Terra Infirma: Geography's Visual Culture*. Routledge, London
- Rose G (1993) *Feminism & Geography: The Limits of Geographical Knowledge*. University of Minnesota Press, Minneapolis
- Saint-Amour PK (2014) Photomosaics: Mapping the Front, Mapping the City. In: Adey P, Whitehead M, Williams A (eds) *From Above. War, Violence and Verticality*. Hurst, London, pp 119–142
- Saint-Amour PK (2011) Applied modernism military and civilian uses of the aerial photomosaic. *Theory Culture Society* 28:241–269. <https://doi.org/10.1177/0263276411423938>
- Saint-Amour PK (2003) Modernist Reconnaissance. *Modernism/modernity* 10(2):349–380. <https://doi.org/10.1353/mod.2003.0047>
- Schuppli S (2012) Impure Matter: A Forensics of WTC Dust. In: Pereira G (ed) *Savage Objects*. Imprensa Nacional Casa da Moeda, Lisbon, pp 120–140
- Schuppli S (2020) *Material Witness: Media, Forensics*. The MIT Press, Cambridge
- Scott JC (1999) *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press, New Haven
- Siebert B (2011) The map is the territory. *Rad Philos* 169:13–16
- Siebert B (2015) *Cultural Techniques: Grids, Filters, Doors, and Other Articulations of the Real*. Fordham University Press, New York
- Smyth P, Fayyad UM, Burl MC, Perona P, Baldi P (1995) Inferring ground truth from subjective labelling of venus images. In: Tesauro G, Touretzky DS, Leen TK (eds) *Advances in Neural Information Processing Systems 7*. The MIT Press, Cambridge MA, pp 1085–1092
- St. Joseph JK, (1945) Air Photography and Archaeology. *Geograph J* 105:47–59. <https://doi.org/10.2307/1789545>
- Steiner F, Fleming B (2019) Design With Nature at 50: its enduring significance to socio-ecological practice and research in the twenty-first century. *Socio Ecol Pract Res* 1:173–177. <https://doi.org/10.1007/s42532-019-00035-1>
- Steyerl H (2013) How not to be seen: A fucking didactic educational. MOV File
- Strauß S (2018) From big data to deep learning: a leap towards strong AI or 'intelligentia obscura'? *Big Data Cogn Comp*. <https://doi.org/10.3390/bdcc2030016>
- Thrift N (2008) *Non-Representational Theory: Space, Politics, Affect*. Routledge
- Veregin H (1994) *Computer innovation and adoption in geography. A critique of conventional technological*

- Models. In Pickles J (ed) *Ground Truth: The Social Implications of Geographic Information Systems*. The Guilford Press, New York, pp 88–112
- Virilio P (1994) *The Vision Machine*. Indiana University Press, Bloomington
- Weizman E (2017) *Forensic Architecture: Violence at the Threshold of Detectability*. Zone Books, New York
- Weyand T, Kostrikov I, Philbin J (2016) PlaNet - Photo Geolocation with Convolutional Neural Networks. *Lect Notes Comput Sci*. https://doi.org/10.1007/978-3-319-46484-8_3
- Xu C (2018) Deep learning and fake geography: creating satellite datasets with Generative Adversarial Networks. AAG Annual Meeting 2018
- Xu C, Zhao B (2018) Satellite image spoofing: creating remote sensing dataset with generative adversarial networks. *GIScience*. <https://doi.org/10.4230/LIPIcs.GISCIENCE.2018.67>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.