Living on Digital Flatlands: Assemblies of Computer Vision

Alex Reid
University at Buffalo

Abstract: This article explores computer vision not only in terms of its use with "intelligent" machines, such as autonomous vehicles, but also more broadly in terms of quotidian, digitally-mediated human vision, such as that experienced through smartphones. Drawing on assemblage theory and radical media archeology, I describe the technological processes that link together materially different spatiotemporal assemblages to produce a posthuman condition of computer vision typified by what Wolfgang Ernst terms the "epistemogenic momentum" of "techno-mathematical configurations." Manuel DeLanda's flat ontologies prove useful here for describing how assemblages operating in different space-times interact. I the turn to Deleuze and Guatarri's techno-semiological stratum, specifically their discussion of the stratum's spatiotemporal superlinearity, to articulate an emerging digital-calculative layer. I argue that understanding computer vision assemblages requires more than setting technical descriptions alongside cultural-ideological critique; it also requires an integrated analysis of the roles digital nonhumans perform in the production of a new visual regime with new capacities and desires. The shared human-nonhuman capacity of computer vision moves sight from the biological and linguistic assemblages of humans into digital assemblages where vision becomes calculable and subject to algorithmic modifications. The results are not only digitally-composed, screened images, but also new means of production, organization, and identity.

In September 2021, the *New York Times* reported on the assassination of Iran's top nuclear scientist, Mohsen Fakhrizadeh (Bergman and Fassihi). T killing was unusual in its reliance upon "the debut test of a high-tech, computerized sharpshooter kitted out with artificial intelligence and multiple-camera eyes, operated via satellite and capable of firing 600 rounds a minute." Fakhrizadeh's killing demonstrates the literal cutting edge of comput vision. As has long been the case in understanding the extent of military-industrial and military-entertainment complexes, we are left to consider the connection between this capacity for remote action and more quotidian devices of computer vision, interaction, and display. In addition to the moral ethical, or political judgments we might make about this particular assassination, the action outlines some of the contemporary pragmatic limits of the spaces in which humans and machines can see or act. Tomorrow or next year, the specifics of those limits will surely be different.

For more than sixty years, the field of computer vision has sought to model human sight, to teach computers to see as humans do, to augment hum vision, and ultimately to exceed the visual limits of humans. From a posthuman perspective, computer vision could also be described as the development of a shared human-nonhuman visual capacity. That is, we are co-participants in the assembly of a kind of vision that incorporates materiality, biology, culture, coding and calculation (these five things being always-already intertwined). These assemblages permeate our lives in the quotidian selfie and online video; military, police, and surveillance technologies; digital art and film; industrial and scientific applications; ethical dilemmas of deepfakes and other fake news; and emerging technologies from AR and VR to autonomous vehicles and robots. As I explore in this essay, the human-nonhuman assemblage of computer vision is more than a continuation of twentieth-century information society or what Lev Manovich terms our "software culture" (38). The shared human-nonhuman capacity of computer vision moves sight from the biological and linguistic assemblages of humans into digital assemblages where vision becomes calculable and subject to algorithmic modifications. The results are not only digitally-composed, screened images, but also new means of production, organization, and identity.

We generally recognize that human vision constructs space and time as much as it represents an independent reality and that such constructions involve desire and ideology. We recognize that technologies of vision also construct space and time. While vision of any kind begins with the input o electromagnetic (light) waves, the subsequent spatiotemporal territory of what is perceived might be assembled from a range of disparate materials with very different results. If, as Walter Benjamin notes, the film camera introduced us to "unconscious optics,"[1] then how might we describe the results of computer vision? (237) Clearly the spatiotemporal scale of calculations running across circuits on a motherboard is quite different from the scale of conscious human time and space. However, the space-time of digital technologies is not simply smaller or faster than biological space-time Computer vision produces new spatiotemporal relations in which we participate just as surely as the telescope shifted our relationship with the stars But how might these new relations be characterized? By combining radical media archeology and assemblage theory I connect what Wolfgang Ern terms the "epistemogenic momentum" of digital technologies with Manuel DeLanda's flat ontological account of how assemblages functioning at different spatiotemporal scales intersect to describe these technical processes and their intersections with human users.

While computer vision technologies can be found in a myriad of specialized applications, they are also as commonplace as the smartphone. With nearly four billion users worldwide, smartphones drive the leading edge of consumer technologies. Projected capacitive touchscreens, OLED display CMOS camera light sensors, lidar, and integrated graphics processors are only some of the computer-visual elements of the latest smartphones. Examinations of specific specialized applications would reveal further dimensions of computer vision but, in my view, the smartphone demonstrates balance between the outlier capacities of computer vision and the material demands of delivering those capacities to billions of human users (or at least such as that balance is optimized by capitalist objectives). In short, the smartphone is the interface through which most humans experience computer vision. While I take into account the specific technical operation of smartphone components, my interest is in the assemblages formed between smartphones and humans in the formation of a shared capacity for computer vision. Ultimately, that shared capacity passes through the smartphone's flatscreen, which is where the technical productions of the smartphone encounter human users.

As I explore here, the smartphone flatscreen is a technical realization of DeLanda's flat ontologies, much in the same way (and with similar caveats a limitations) as hypertext realized the rhizome. Flat ontology describes a non-hierarchical, non-deterministic relationship between two or more assemblages operating at different spatiotemporal scales and serves as an intersection through which forces and information may travel. That is, the flatscreen provides bi-directional feedback for both human and machine without which assemblages of computer vision could not form. The digital, technological processes that produce flat screen images present a new articulation of time and history that diverges from the continuous motion and measurement that characterizes analog timekeeping as well as historiography. Among these assemblages are the human users and the communitie we form through and with our screens.

While these methods produce useful insights, common concerns with flat ontologies suggest that they erase political power relations and ignore history. Rosi Braidoitti writes, "The political strength of the relational interconnection among different entities in the world is disavowed by the flat ontology espoused by Latour and his object-ontologist champions (Harman 2014)… no amount of claim to the equality between human and non-human actors, which ANT has voiced so explicitly, can compensate for the lack of an epistemology that does justice to the power structures of contemporary subjects" (56).[2] The association of Latour with flat ontology has always struck me as strange. I know of no instance in which he uses the term. His argument for symmetry, which might be misread as similar, is nothing like this. As Latour writes, "ANT is not, I repeat is not, the establishment of some absurd 'symmetry between humans and non-humans'. To be symmetric, for us, simply means not to impose *a priori* some spurious asymmetry among human intentional action and a material world of causal relations" (76). Similarly, from the object-oriented ontological perspective (where the term is used though not necessarily espoused without critique), flat ontology asserts "*all things equally exist, yet they do not exist equally*" (Bogost 11; emphasis in original). As I will discuss, both the ANT and OOO positions bear only modest connections with DeLanda's fla ontology, but none of these positions assert that there are no material-historical differences among humans and nonhumans. To the contrary, particularly in DeLanda's and Latour's cases, they argue that *all* of the differences among humans and nonhumans are material and historical rather than being created *a priori* by some divine creator who built essential hierarchies outside the realms of space and time (which, as far I can figure, is t only alternative to asserting these differences are material and historical).

That said, as I discuss here, new materialism's radical decentering of humans can make it difficult to reconnect with human concerns. In the case of DeLanda, this difficulty arises from his inattention to desiring-production. Desiring-production, while not anthropocentric in itself, names the force animating assemblages (at least for Deleuze and Guattari). Espousing this perspective, Braidotti identifies the necessity of a vibrant force that is "driven by the ontological desire for the expression of its innermost freedom" (47). However, this desire is *not* anthropocentric. Instead, it arises in th relations among humans and nonhumans: "[t]his understanding of matter animates the composition of posthuman subjects of knowledge as embedded, embodied, and yet flowing in a web of relations with human and non-human others. Posthuman subjectivity is an ensemble imposed of *zoe*-logical, geological, and technological organisms--it is a *zoe*/geo/techno assemblage" (47). In this *zoe*/geo/techno assemblage, we can see Deleuze and Guattari's strata, all of which are activated by desiring-production, adapted in Braidotti's new materialism/posthumanism. Taken togeth adding these strata and desiring-production to flat ontological and radical media archeological methods makes it easier to address the political and ethical concerns of a shared human-nonhuman computer vision without reasserting anthropocentrism.

The Space-Time of Flat Ontologies

The origin of flat ontology is easy to trace to Manuel DeLanda's *Intensive Science and Virtual Philosophy,* though its usage after that becomes more complicated.[3] DeLanda deploys the term in his exploration of biological individuation. While Darwin provides a historical materialist starting point fo evolution, DeLanda notes "the idea that *species are individuals*, not kinds, has only recently (and still controversially) gained ground" (*Intensive* 46).[4 Unlike the conventional typological explanation in which I am an individual, which is a kind of *homo sapiens sapiens*, which is a kind of hominid, etc. each of these "kinds" is instead another individual, just at a different spatiotemporal scale. DeLanda writes, "one philosophical consequence of this new conception of species must be emphasized: while an ontology based on relations between general types and particular instances is *hierarchica* each level representing a different ontological category (organism, species, genera), an approach in terms of interacting parts and emergent wholes leads to a *flat ontology*, one made exclusively of unique, singular individuals, different in spatio-temporal scale but not in ontological status" (47). In other words, each individual human is part of a species, which is also an individual with emergent qualities that differ from those of the individual pa

(i.e., actual humans) that comprise it. That is, to put it bluntly, the species has existed for millennia and spans the world, while I do not. This biologic example of the non-hierarchical (i.e., flat) relationships among assemblages operating at different spatiotemporal scales serves as a useful conceptu tool for approaching the relations among other assemblages with spatiotemporal differences.

In the absence of essential, hierarchical relations, flat ontologies (and the broader assemblage theory of which they are a part) investigate the spatiotemporal relations among assemblages in terms of material history. That is, to continue with the biological example, if my individual humannes is not an essence granted to me by my situation inside my species, which gains its essential qualities from "natural law," a divinity, or something similar, then the biological relationship among me, other humans, and our species can *only* be described as historical. A cursory examination of the basic topics or even the titles of DeLanda's works, such as *War in the Age of Intelligent Machines, A Thousand Years of Nonlinear History*, and *A Ne Philosophy of Society* would reveal an explicit interest in social and political subjects. Indeed, a significant part of the work that flat ontology and assemblage theory does is to account for the material-historical emergence and operation of socio-cultural power relations and economic structure. Nonlinear dynamics are not only integral to DeLanda's theory of history but to new materialism in general. Similar to flat ontologies, nonlinear histor describes assemblages operating at different spatiotemporal scales and interacting in non-hierarchical ways. For example, geological and climatological spatiotemporal scales, prior to the Anthropocene at least, were considered to operate beyond the ability of social assemblages to affect, though clearly, they affected us. However, the relationship among every coal fire, smoke stack, car engine and the global climate, with the subsequent impact of climate change on a range of human activities from farms to coastal cities, is an all-too-familiar example of nonlinear historica processes at work.

The nonlinear material history of climate change is more than an apt example of the concept; it is an important intersection between computer vision and political-ethical concerns. Contemporary climate science is impossible without the computer vision of satellite imagery, digital models and simulations. If, as Bradiotti writes, environmental humanities, along with digital humanities/media studies are central matters of the "critical posthumanities," then surely their intersection is crucial (116). Jussi Parikka's *Geology of Media* is one of the more recognized examples of that intersection. As he describes, those connections were latent in Foucault's interest in an archeology of knowledge, which inspired Kittler and media archeology, and in Deleuze and Guattari's geological assemblage theory. I am following a similar vein here and asserting that flat ontologies map connections across these diverse spatiotemporal assemblages from cosmological and geological strata to the bio/*zoe*-logical stratum and the techn semiological stratum of digital hardware and global data networks. It is not sufficient to examine any of these strata on their own. Parikka describes these connections in terms of medianatures, "a regime constituted as much by the work of micro-organisms, chemical components, minerals, and metals as by the work of underpaid laborers in mines or in high-tech entertainment device component production factories, or people in Pakistan ar China sacrificing their health for scraps of leftover electronics" (14). Medianature, much like Donna Haraway's natureculture, is a concept that becomes increasingly useful when used to investigate particular contexts, such as the computer vision of smartphone user assemblages.

These new materialist, media archeological investigations can be complementary to, rather than in opposition to, other methods. As Ernst contends "While a more discourse-oriented analysis (like Science and Technology Studies) performs a critical anatomy of the underlying cultural, ethical, political, and economic biases which are at work in algorithms, radical media archaeology, in alliance with object-oriented ontology, focuses on the epistemogenic momentum which it derives from within techno-mathematical configurations" ("Existing" 4). Pointedly, if we are studying how media technologies operate, we have to study their operation. Ernst leaves open the question of how these archeological and discursive analyses intersec As I discuss here, one possible approach lies with assemblage theory and its flat ontological capacity to describe the intersections among assemblages operating at different spatiotemporal scales without insisting on a particular hierarchy. That is, the algorithmic operation of computatio hardware at spatiotemporal microscales intersects with the operation of larger institutional assemblages without one determining the other.

Of course, space and time are themselves part of these assemblages. That is, assemblages construct their own space and time in manners particul to their material histories. This notion may be counterintuitive in that even though spatiotemporal scales range from the subatomic to light years, we tend to view them all as part of a common system of measurement. However, those measurements are only another assemblage. It is not only that t space-time of the Earth or the solar system are immense compared to human scales. Their mechanisms are different. If the spin and orbit of planets around the sun and the sun around a galactic core are produced by gravity, then our days and years intersect those assemblages. However, our calendars are not produced by gravity. They are produced by observations of the sun in the sky and an increasing ability to observe and calculate, which, of course, eventually shifts us to a heliocentric model. As I explore here, digital media construct space and time as calculable dimensions, opening both to algorithmic manipulations. Our taps and clicks become part of an oscillating chorus of screen refresh rates, CPU cycles, and netwo speeds. In our digital-cultural participation in posthuman medianatures ranging from the cosmological and geological to the digital and technologica we are not simply in-between, in scales ranging from bigger and slower to smaller and faster; we are engaging in ontologically different spatiotempo assemblages that do not nest inside one another.

Speculating on Smartphone Flats

To investigate the medianatural construction of algorithmic optics, and particularly the role of flat screens in it, I turn first to their technical operation. In addition to the displays themselves, the surface of the smartphone relies on two other late twentieth-century technical innovations: touch screen technology and glass-ceramic. Glass-ceramic provides an impact-resistant surface. Contemporary glass-ceramics combine the amorphous solid of glass with a controlled crystallization of ceramics, and smartphones are only one of many military, medical, industrial, and commercial applications of the technology. Without the durability of this surface, the practical mobility and usability of these devices would be severely limited. Beneath this protective layer is organic light emitting diode technology (an OLED display) as well as projected capacitive touch technology. In turn, there is the smartphone's hardware that interacts with the screen and an infrastructure of power and data networks from the "last mile" reaching into our homes and pockets to undersea data cables and servers in temperature-controlled vaults. These assemblages are smeared across the surface of the planet and reach out to satellites in low-orbit (for GPS). All of this technology has become mundane to smartphone users, even though the technologies that comprise the contemporary mobile internet only became assembled in the last 15 years. Alternately, following on Jussi Parikka's geology of media, we can trace the history of the strata and assemblages that produce the more than sixty metals found in a smartphone and Deleuze and Guattari's phylogenetic lines that "travel long distances between assemblages of various ages and cultures (from the blowgun to the cannon? from the prayer wheel to the propeller? from the pot to the motor?)" (407). Afterall, as early as the Bronze Age, we can find the beginnings of written codes, metallurgy, ceramics, and glassmaking that will come together to form the smartphone millennia later.

These are some of the assemblages that intersect with the population of contemporary smartphones, and each offers an opportunity for tracing the emergence of the particular tendencies and capacities these technologies exhibit. At the same time, the smartphone assemblage exhibits its own emergent spatiotemporal existence. Unlike mechanical clocks that measure time by a wound spring, computer hardware constructs time through the piezoelectric punctuations of a crystal oscillator. As Wolfgang Ernst writes, "media archaeology principally reminds us of the rigid frequency regime of digital computing in its electronic instantiations, which is counting rather than timing. Digital computing treats time as discrete, thereby transforming time itself into computation" ("Existing" 14). He continues, "different from the phenomenological time of lived experience, data processing is no continuous 'stream' at all, but nonlinear microstorage and transfer" (15). This transformation of time into computational nonlinear microstorage is comparable to the more familiar digitization of text and other analog media. There too we see the shift from the continuous measurement and variation of a sound or film recording or handwriting to a series of discrete, sampled points that are then made available for computational manipulation.

In terms of flat screens, this discretization of time is measured by its refresh rate. For contemporary smartphones this is 60-120Hz (i.e., 60-120 times per second) but will soon be as fast as 240Hz (rivaling high-end computer monitors). With each refresh, each pixel is charged to emit a color. The pixel neither receives nor requires information regarding whether it is part of a photo, a video, a background screen, an icon, an animation, a text document, or whatever. In fact, pixels do not receive coded messages of any kind but rather an electrical charge that produces a color. These pixels are composed of blue, red, and green subpixels arranged in a diamond, each receiving separate charges. Working backwards, those electrical pulses are sent to the subpixels by graphics processor units integrated into smartphones, and those graphics processors receive their instructions from CPUs that must handle various inputs, each made calculable by the oscillations of a crystal. Has the touchscreen been activated? Is the accelerometer initiating a shift from portrait to landscape? Has a notification arrived via the network that needs to be displayed? But none of these questions matter to the screen itself. So, at minimum, there are three assemblages intersecting here: the processor with its binary strings, the screen with its charged pixels, and the human watching the screen. While neither the processor nor the human may give much significance to a series of charged pixels, through interface of those pixels, the two enter a shared visual regime. For that to happen however, at least on a smartphone, touchscreen technology must also be employed.

Projected capacitive touch technologies introduce another spatiotemporal calculation onto the surface of the smartphone. Where the screen image is produced by millions of electrical pulses counted out ~60 times per second, our touches are also electrical signals that are similarly captured within the phone's calculative spatiotemporality. Without those gestural inputs, there is very little for the phone to do except wait and remain alert for messages from the data network. The touchscreen is constructed from two layers, columns and rows, of capacitive sensors. A low-level charge runs continuously through the screen while the phone is operating. When users touch their screens with one or more fingers, that charge is disrupted. This technology made contemporary smartphones and tablets possible, as without it users would require the miniature keyboards familiar from Blackberry phones. Contemporary touch technologies provide additional interaction. These touches activate algorithms: opening apps, following links, typing letters, liking posts, making purchases, etc. In short, touch technologies turn the screen into a user-programmable surface. These touches also enter into the calculative temporality of digital media. This calculative quality is likely most clear in video game play, where the timing of touches is often crucial. However, every instance of touch occurs within a computational cadence and in a computed, screen location. As users click and drag a virtual slider to adjust the quality of an image (saturation level, for example), they respond to the analog shifts of color light projected from the screen, but the phone is making discrete, incremental, and mathematical alterations in the instructions it is sending the graphics processor, which in turn is varying the electrical charges sent to millions of pixels. Users see the light waves that strike their eyes; touch screens register the shifting of current across conductive layers; the phone's programming and processors turn that information into adjustments to an image file in RAM and send instructions to the screen, and the screen changes its display.

Somewhere in that distributed process, the capacity to see an image emerges, and often that distributed process involves actors beyond the user a a single device. While the pixel knows nothing of the image in which it participates, I can recognize my face in a selfie as different from the background and so can my photo editing software. The social media platforms where I post my selfie see something similar to what my image editir software sees. They invite me to confirm that the box they have drawn around part of an image is the face of a friend or relative.[5] Is the identificatio of a face in a digital picture a hallucination? If so, it's one that we are sharing with software. On the other hand, if I see that image as a flat area of colored but otherwise undifferentiated pixels, is that a hallucination? I suppose we might ask an AI Magritte, who might tell us that representations have always worked this way and that, in some respects, pixels are not different from chemical exposures or paint strokes. In other respects, of course, pixels are quite different as they are calculated. As I have been describing, pixels are the technical targets of image processing algorithms. Whether those algorithms are the compression tools that create GIF and JPG images or generative adversarial networks (GANs), which are employe to produce deep fakes, they write instructions for pixels. At the same time, the human response to these pixelated images provides valuable feedba to those adaptive algorithms. We are co-creating a shared assemblage of computer vision. Without these technologies, our capacities to see and touch the world are severely curtailed. Just as certainly, however, our smartphones have little opportunity to see and interact with the world without our participation. Admittedly, only anthropomorphism would allow us to assert that computers see or act. But as Jane Bennett argues,

> an anthropomorphic element in perception can uncover a whole world of resonances and resemblances-sounds and sights that echo and bounce far more than would be possible were the universe to have a hierarchical structure... [and] catalyze a sensibility that finds a world filled not with ontologically distinct categories of beings (subjects and objects) but with variously composed materialities that form confederations. (99)

In this context, the anthropomorphism of computer vision allows us to trace the passage of affects and effects across the flat ontological connectior between human users and their computers. For Bennett, drawing on Latour, the representation of nonhumans is both a scientific-empirical challenge and a political-cultural one, and anthropomorphism becomes a strategy for recognizing isomorphisms across those materially and historically managed divides. As she writes, "Clusters of neurons in a human brain, groupings of buildings in a city, and colonies of slime molds all have been shown to follow similar organizational rules" (100). In computer vision, "vision" is already an anthropomorphic gesture as well as one that recognizes an isomorphism. Computers do not "see" as we do, but they are responding to the same input of light waves, and, because computers and humans have developed together, we form an assemblage (though Bennett prefers "confederation" here) that produces representations of an empirically observed world but simultaneously produces political representations in the manner those observations are necessarily valued and made available f calculation.

Mining the digital stratum

While we often move from a camera's representation of the world to political representations and judgements, as when a defendant is caught on camera committing a crime, in philosophical terms there are some missing steps. In *Alien Phenomenology*, Ian Bogost discusses our use of metaphors in understanding nonhuman experiences. He writes that linking camera sensors to voltages to processors to graphics controllers to scre pixels "involves phenomenal daisy chains, built of speculations upon speculations as we seep farther and farther into the weird relations among objects. The philosophical effort to bind such metaphors is nontrivial, amounting to a complex lattice of sensual object relations, each carrying an inherited yet weaker form of metaphor with which to render its neighbor" (81). As I have been describing the shift from one assemblage to the next i linear fashion, a daisy chain is an appropriate description. However, any given assemblage has multiple intersections with others and any of those paths might be followed. So, I can follow a daisy chain from the glass ceramic surface of a smartphone to the touchscreen and OLED layers below i or I can follow the manufacturing process or the chemical properties of the molecules or the history of ceramics or material sciences. Bogost's gene point still holds, however. There's a non-deterministic, non-hierarchical, flat-ontological relation among these assemblages and daisy chains identify the philosophical challenge of moving across assemblages.

Though DeLanda takes the decision to fold Deleuze and Guattari's concept of strata into his assemblage theory, strata are conceptually useful for addressing this challenge.[6] Deleuze and Guattari describe three main strata: the geological, the biological, and the anthropomorphic (though following on Ian Buchanan, I prefer "techno-semiological" for the third strata) (Buchanan 28). Among the strata, there is a linear historical progressio but each stratum also affects and creates conditions for the others through nondeterministic relationships. As Buchanan writes, "'we' humans deper on the properties of the earth for our existence (geology) and 'we' depend on the properties of our bodies for what 'we' can do on the earth (biology but 'we' constantly exceed those limits in the outpourings of our minds… the production of signs (both symbols and language) enables the third stratum to translate the other two and in a sense range beyond them" (29). However, as the scare quotes around "we" suggest, there's more to this matter, as has been the subject of this essay: it is not simply or solely "we" who are exceeding these limits, but rather humans assembled with nonhumans. Deleuze and Guattari recognize this, writing, "Content should be understood not simply as the hand and tools but as a technical social machine that preexists them… Expression should be understood not simply as the face and language, or individual languages, but as a semiotic collective machine that preexists them" (63). With this in mind, the techno-semiological stratum should not be seen as essentially or exclusively *of* o *for* humans, though we clearly participate in it, as we do in the geological and the biological strata. From clay tablets to rare earth minerals in smartphones, the geological stratum has had a material role in the techno-semiological, as has the biological, since techno-semiological processes

have had to interface with human biology: sounds our voices could make and ears could hear; signs, gestures, and images our eyes could see and hands and bodies could produce. With digital technology, cosmological strata must also be considered, as the current boundaries of the techno-semiological push up against the properties of light, subatomic particles, and quantum states.

Despite the constraints the cosmological, geological, and biological put upon the content and expression of the techno-semiological stratum, Deleuze and Guattari identify the stratum's unique capacity in its temporality: "*Vocal signs have a temporal linearity, and it is this superlinearity* that constitutes their specific deterritorialization and differentiates them from genetic linearity" (62, emphasis in original). Following their example, when humans develop speech, they move from the biological strata of animal noises that exist within a fixed temporality into an ability to reference other times and spaces. This capacity for spatiotemporal translation marks the techno-semiological stratum and represents "the ability of language, with its own givens on its own stratum, to represent all the other strata and thus achieve a scientific conception of the world" (62). This act of representation is another link in the daisy chain of strata and assemblages. However, the spatiotemporal superlinearity of techno-semiotic codes permits a new kind of flexibility and reversibility. For example, this sentence can be revised, moved or deleted. As Buchanan observes, "the meaning of a phrase or sentence cannot be arrived at additively… We may be able to guess where a sentence is going--as autocorrect and autofill on our smart phones try to do, with varying degrees of success--but even allowing for the built-in redundancies of syntax and grammar we cannot know with absolute certainty where it will end up" (42). The smartphone is an interesting example, and no doubt accurate (though I will confess that often I am not sure exactly where or how my own sentences will end, until they do). However, the smartphone also points to an emergent digital layer, which, as Ernst contends, introduces a new calculated temporality. While machines may be no better than humans at finishing each other's sentences, increasingly machines are able to write sensible scripts of their own. They are able to converse. They identify us and the world we share and produce meaningful sounds, voices, images and videos, including algorithmically generated fictions such as deepfakes. Hypothetically any image, audio, or video can be made to look and sound like anything else. It is this capacity to turn time and space into calculable territories subject to mathematical interventions that characterizes the new, expanded superlinear spatiotemporality of the digital-calculative stratum or layer.[7]

Furthermore, the digital-calculative layer introduces a posthuman condition in which humans are decentered. From the inception of symbolic behavior through to the twentieth century, it was possible to argue that technical, social and semiotic collective machines required direct human cognition to operate. Books could neither write nor read themselves or one another. As Deleuze and Guattari write, "It's all in the head… What we are trying to say is that there is indeed one exterior milieu for the entire stratum, permeating the entire stratum: the cerebral-nervous milieu" (64). However, with the arrival of the digital-calculative layer, it is at least possible to consider that this containment within the "cerebral-nervous milieu" is no longer the case. While we have not (and may never) reach the point of artificial general intelligence, the posthuman era has been shaped by our confrontation with increasingly "smart" machines that operate with greater independence from humans than earlier technologies. This digital posthuman condition is familiar to anyone who has taken driving directions or book recommendations from an algorithm. Indeed, the difference between a printed road map and digital assistants directing drivers when and where to turn aptly captures the shifts arriving from our digital lives. We remain in the loop with these technologies but our interactions with them introduce us to a new space and time.

Our experience of the digital-calculative layer points to another way Deleuze and Guattari describe strata. As Buchanan observes, strata help us "analyze and explain the fact that everyday life is experienced by most people as multi-layered, without necessarily being organized and interconnected" (48). For example, in our work "we engage with machines, we engage with people, we engage with command structures and schedules, and a whole host of other kinds of inputs and variables" (48). He turns to the arrival of digital technology and its effects on academic life, adding that "The more interesting and much less maudlin response, though, would be to ask whether any new assemblages have appeared. Is a lecture online different in kind to one viewed 'live' in a lecture theatre? Is an e-book different in kind to one made from trees?" (49). More generally, most people would say that they have a digital life, a new layer, that intersects the rest of their lives. They have digital identities and online communities and friends. They engage in novel semiotic practices that include not only emerging genres and media formats but also an entirely new form of meaning-making arising from the algorithmic analysis of big data to which they each contribute. Conceived this way, the digital-calculative layer clearly has its own technical social and semiotic collective machines. Before, as Deleuze and Guattari did, we could only speak of techno-semiological representations of other strata. Books, printing presses, radio stations, and television sets could not act on their own. However, as we move from the late twentieth-century third industrial revolution (information) into a fourth industrial revolution (characterized by smart machines and automation), the assemblages formed among humans and digital technologies connect with this layer to establish a new superlinear spatiotemporality in which the calculated codes do not simply represent the other strata but also produce a new materiality. In our daily lives this new layer is apparent not only in the always-on connections and availability of media and data and our digitally-mediated social relations but also in the developing algorithmic regime organizing our lives and the broader social order. Returning to the daisy chain of assemblages in smartphones, from the input of light into the camera lens to the output of light from the flatscreen, there are a series of assemblages that connect across strata moving through the various parts of the device: cosmological light waves, radio waves, and electrons; geological lenses, ceramics, and metals; biological organic molecules; and techno-semiological codes. No one stratum supersedes the others, but on the digital layer, a superlinear calculated temporality codes a new visual regime punctuated by network speeds, clock cycles and refresh rates.

As significant as the emergence of a digital-calculative layer might be, this technological accounting of these assemblages is incomplete without consideration of the role of desiring production. Buchanan reminds us that assemblages are ultimately about desiring-production: "*It is desire that selects materials and gives them the properties that they have in the assemblage*. This is because desire itself is productive" (emphasis in original, 5(So these assemblages should not be confused with "merely" extremely complex technical devices. If desire is "defined as a process of production without reference to any exterior agency," as Deleuze and Guattari write, then it is not necessarily human (154). Specifically, as they also write, "the rationality, the efficiency, of an assemblage, does not exist without the passions the assemblage brings into play, without the desires that constitute as much as it constitutes them" (399). As such, to conceive of computer vision as an assemblage is to describe not only a complex device but a productive machine that both assembles, and is assembled by, desire. As elusive as the concept of desire might be, without it, there is a danger for the new materialist analysis of digital media to become little more than an imprecise description of technical functions better understood by engineer Radical media archeology provides part of a response to this problem. Ernst writes, "There is no technology that is purely 'digital,' material- and energy-free information in essence – just like the monetary value is always materialized, be it in the coin, the bank-note or its electronic equivalent in bitcoin cryptocurrency. Matter is assigned a different agency when it becomes mathematically informatized instead of simply technically formatted" (*Techo*logos 3). In digital media assemblages the contents and expressions of the techno-semiological stratum intersect the geological stratum whe they are etched into the surface of silicon oxide, "as electronically programmable matter" (2). In this etching, desiring-production is operating. As wit monetary currency, it is easy enough to see familiar human desires in the creation and use of digital media, but, as with money, it is also easy enoug to see desires flowing through the assemblage rather than anthropocentrically imagining humans as their alpha and omega. In short, the re-incorporation of strata and desiring-production into assemblage theory provides a transition between the focus of a new materialist, radical media archeology on the epistemogenesis of techno-mathematical configurations within digital devices and the ethico-political concerns Ernst characterize as the focus of "discourse-oriented analysis." It is, undoubtedly, a transition rather than a synthesis, a means for investigating the non-hierarchical, non-deterministic, flat ontological encounters in this daisy chain of beings.

In short, the reintegration of strata and desiring-production into assemblage theory provides a more thorough means of investigating the human-nonhuman link in computer vision than that provided by the daisy chain of metaphorical anthropomorphizing. It brings the geology of media from technologies' beginnings and endings as raw materials and later as e-waste into the continual looping of the assemblage where cosmological and geological strata form the piezoelectric pulses that are made calculable in the digital layer. The camera lens, like the human eye, senses light, but it cannot see without some intent, some desire, neither can it see without establishing some relation, some stimulus or intensification. That is, to see is to see something, even if it is to see "nothing," which creates its own stimulation. In doing so the assemblage of computer vision loops through the biological strata and human sight. In theory, the aim of artificial general intelligence may be to remove human deliberation and action from the loop automated processes. In the case of autonomous vehicles, from a certain perspective this is accomplished; the human need not touch the steering wheel. Of course, the autonomous vehicle does not choose its destination. Similarly, from a certain perspective, we say the human passenger chooses the destination, but in what sense do we choose to commute to work, run shopping errands on the weekend, or even take a relaxing drive through the countryside? Here too we are part of assemblages looping through strata and participating in desiring production. In the presumed near future of everyday autonomous vehicles, the daily commute becomes additional time for work productivity, shopping trips blend with catching up or social media, and highway drives become opportunities for watching the newest blockbuster movie. In other words, the former drivers now do what the other passengers in their cars have been doing for a decade. Of course, for all those other passengers those interactions have occurred through their smartphones.

Fakhrizadeh's assassination provides a starker distinction among competing human and nonhuman desires and the role intelligent machines might play. As DeLanda writes, "In adversarial situations, being forced to treat a machine as an intentional system may be considered as a good criterion f mechanical intelligence. In the case of predatory machines, not only would we have to fight them on the 'intentional plane' but we may also assume that they would treat us, their prey, as predictable assemblages of beliefs and desires. It would be, then, a clash of 'minds' or of 'rational wills'" (*War* 157). The state's desire to see and act on a global scale shapes the development of computer vision as surely as the corporate desire for robotic factories, automated vehicles and delivery drones: such is the long story of the military-industrial complex. It is another reminder of Walter Benjamin warning that "the destructiveness of war furnishes proof that society has not been mature enough to incorporate technology as its organ, that technology has not been sufficiently developed to cope with the elemental forces of society" (242). At the same time, the material and techno-semiological capacities of digital media have their own expressions. That is, while the state may desire the capacity for instantaneous, error-free, global surveillance and action, there are material limits to digital mediation. As the potential target of a drone strike or robot assassination, respondir to the all-too-human desires of distant military technicians or their political handlers is less useful than predicting the intentions and capabilities of th "predatory machine." Those predictions require understanding the calculative space-time of digital machines and the capacities they engender for a broader assemblage of computer vision. Our quotidian encounters with consumer computer vision and intelligence are less adversarial and perhaps even symbiotic, but, as we negotiate our participation in these assemblages, we might also benefit from understanding how assemblages of compu vision draw us into new desires and spatiotemporal relations with the world.

Works Cited

Benjamin, Walter. "The Work of Art in the Age of Mechanical Reproduction." Translated by Harry Zohn, *Illuminations,* Harcourt, Brace & World, 1968, pp. 217-42.

Bennett, Jane. *Vibrant Matter.* Duke UP, 2010.

Bergman, Ronen and Farnaz Fassihi. "The Scientist and the A.I.-Assisted, Remote-Control Killing Machine." *The New York Times,* 24 September 20: nytimes.com/2021/09/18/world/middleeast/iran-nuclear-fakhrizadeh-assassination-israel.htm.

Bogost, Ian. *Alien Phenomenology, or What It's Like to Be a Thing.* U of Minnesota P., 2012.

Braidotti, Rosi. *Posthuman Knowledge*. Polity, 2019.

Buchanan, Ian. *Assemblage Theory and Method.* Bloomsbury, 2021.

DeLanda, Manuel. *Assemblage Theory*. Edinburgh UP, 2016.

-----. *Intensive Science and Virtual Philosophy.* Continuum, 2002.

-----. "Ontological commitments." *Speculations: A Journal of Speculative Realism,* vol. 4, 2013, pp. 71-73.

-----. *A New Philosophy of Society: Assemblage Theory and Social Complexity*. Continuum, 2006.

-----. *A Thousand Years of Nonlinear History.* Zone Books, 1997.

-----. *War in the Age of Intelligent Machines.* Zone Books, 1991.

De Landa, Manuel, and Graham Harman. *The Rise of Realism*. Polity, 2017.

Deleuze, Gilles, and Félix Guattari. *A Thousand Plateaus: Capitalism and Schizophrenia*. Translated by Brian Massumi. U of Minnesota P, 1987.

Ernst, Wolfgang. "Existing in Discrete States: On the Techno-Aesthetics of Algorithmic Being-in-Time." *Theory, Culture & Society,* Nov. 2020, doi:10.1177/0263276420966396.

-----. *Techno*logos *in Being: Radical Media Archaeology & the Computational Machine*. Bloomsbury, 2021.

Harman, Graham. *The Quadruple Object.* Zero Books, 2011.

Huneman, Philippe, and Denis M. Walsh, eds. *Challenging the Modern Synthesis: Adaptation, Development, and Inheritance*. Oxford UP, 2017.

Manovich, Lev. *Software Takes Command*. Bloomsbury, 2013.

Parikka, Jussi. *A Geology of Media.* U of Minnesota P, 2015.

Notes

1. To put that phrase in greater context, Benjamin writes "the camera introduces us to unconscious optics as does psychoanalysis to unconscio impulses" (237). He arrives at this conclusion, in part, by recognizing that "by exploring the commonplace milieus under the ingenious guidar of the camera, the film, on the one hand, extends our comprehension of the necessities which rule our lives; on other hand, it manages to assure us of an immense and unexpected field of action" (236). It is not simply that the camera shows us the world in a new way, The camera produces a new spatiotemporal territory that produces new knowledge and engenders new agency. Within the camera's territory, human psychology is newly expressed and revealed, but so too are broader political movements from Soviet montage to fascist crowds. For Benjam the former politicized art while the latter rendered politics aesthetic. The camera, though, served both purposes. Here I raise similar questions How might we describe the spatiotemporal territory of computer vision? What new epistemological and ontological (and by extension ethical and political) conditions does it produce for us?

2. Some of those "object-ontologist champions" have previously asserted that the Deleuzian emphasis on immanence that Braidotti prefers, wh

they term "undermining," also creates an undifferentiated, flat ontology, an accusation that Braidotti turns back on them here (see Harman, 7-19). Is either of them right? Does it matter? What appears clear is that philosophers/theorists do not tend to describe their own work as being incapable of recognizing differences in the world or of being of no use to their fellow humans.

3. In his conversation with Graham Harman in *Rise of Realism,* DeLanda says, "I never realized that the expression 'flat ontology' could be used such a variety of different senses. I should be more careful when I use it because, as you point out, there are different ways in which one can flatten ontologies" (87). Here he is referring to different ways to flatten within the discourses of philosophy. For example, the flattening done by empiricism is quite unlike that of DeLanda's realism. However, these variations become more pronounced beyond philosophical discussions of ontology, as when flat ontology intersects with the cultural and political concerns of rhetoric.

4. The modern synthesis of Mendel's inheritance and Darwin's population change forms the basis of a unified theory of evolution in the early and mid-twentieth century. It has successfully adapted over the last century, though "it has become commonplace in the last few years to hear claims that the modern synthesis should be wholly rethought (Laland et al., 2014), revised, or extended (Pigliucci & Müller, 2010)" (Huneman and Walsh, 1). As Huneman and Walsh explain, "the challenges, in one way or another, concern inheritance, development, and adaptation—and the relations between them" (13). DeLanda positions himself with those calling for a more dramatic rethinking of the modern synthesis, and this is reflected in his interest in evolutionary development biology, a field that developed in the 1990s.

5. On November 2nd, 2021, Meta, the parent company of Facebook, announced it would shut down the Facial Recognition system on Facebook, though it plans to continue using the technology and other biometrics as part of its proposed "metaverse" application.

6. I am not suggesting any reconciliation or reformulation between Deleuze and Guattari's strata and assemblages and DeLanda's assemblage theory, though the former explicitly influences the latter. Ian Buchanan specifically objects to DeLanda's approach to assemblages, but I am not interested in entering into that debate either. Instead, I simply see strata as offering a related conceptual approach to the question of how digital-technological assemblages construct spatiotemporal relations and how our intersections with those assemblages establish a computer visual regime.

7. In their discussion of strata, Deleuze and Guattari mention substrata, epistrata, and parastrata to account for the various non-hierarchical relations among and within strata, though ultimately there is little to distinguish among them. "The epistrata and parastrata subdividing a stratum can be considered strata themselves (so the list is never exhaustive)" and "one stratum is always capable of serving as the *substratum* of another, or of colliding with another, independently of any evolutionary order" (502, emphasis in original). I have decided to employ the more generic "layer" (as Deleuze and Guattari also do) as a digital-calculative layer might be termed a stratum (or an epi-, para-, or sub-stratum) at any given instance.

Cite this Essay