

Lee, Rosemary. "Machine Learning and the Mediating Tendencies of the Image". Presented at Dark Eden - The Sixth International Conference on Transdisciplinary Imaging at the Intersections between Art, Science and Culture, University of New South Wales Faculty of Art and Design, Sydney. 2021.

Machine Learning and the Mediating Tendencies of the Image

Rosemary Lee
IT-University of Copenhagen
Copenhagen, Denmark

Abstract

The technological mediation of human perception that occurs through images influences not only how they are produced and experienced but also how they are interpreted. The present incorporation of machine learning (ML) into various forms of visual media offers insight into this issue by enabling images to be produced as the result of the statistical analysis of datasets. Computational relations that are extracted and inferred between features within images help to construct learned representations which are in turn used to generate images. This results in a form of computationally-determined representation that is informed by the interpretive processes performed by machines. This paper addresses several ways in which current notions of image production prove inadequate for the description of the visual artefacts of ML, leaning heavily on historical narratives regarding the technical production of images and even perpetuating inaccuracies. It seeks to clarify the mediating role played by visual technologies and to demonstrate how images produced using ML offer new ways of approaching theories of the image. This investigation considers how the participation of highly technical systems in visual media ultimately contributes to a critical re-evaluation of the image and what this may mean for visual culture.

Introduction

Visual technologies play an important role in the mediation which occurs both in the production and the interpretation of images. Technology's role in expanding human experience and ability can also influence how those are interpreted. This is especially relevant as highly automated visual systems grow in terms of technical capacity but also become increasingly integrated into

many applications. Current forms of algorithmic media, such as images produced using machine learning (ML), demonstrate the interrelation between visual perception and the production and interpretation of images in a particularly relevant way. And as a paradigm of image production, ML raises several important, unresolved questions about the interrelation between human and machine forms of visual processing.

ML refers to both a field of artificial intelligence (AI) research and an approach, “in which machines ‘learn’ from data or their own ‘experiences.’”¹ When applied to visual tasks such as the generation or analysis of images, ML enables complex visual processing tasks to be performed by computers, often in a highly-automated fashion. Familiar applications of ML include facial recognition, influencing the display of online content and the generation of images based on analysis of existing data, facilitating the automated generation of visualisations from data analysis, as well as the commensurate analysis of images.

In addition to becoming increasingly prevalent in visual media, ML has also recently experienced a surge of interest from artists as well as theorists, who have been working with ML in a practical capacity, as well as reflecting critically on its significance to visual culture. The appropriation of ML by artists enables novel aspects of ML to be explored in terms of humanistic discourses, but it also brings with it historically charged ideas concerning the role of technology within art and artistic practice. While the widespread use of ML in visual media is a recent phenomenon, how it is theorised often links current forms of visual media to ongoing narratives about the role of machines in image-making. This has the benefit of contextualising newer forms of media in relation to their “old media”² precursors, but also brings with it several unresolved issues, especially regarding the autonomy of machines from human intentionality or perception.

Methodology

This paper gives an overview of discourse surrounding the generation of images using ML, examining this topic through a survey of artistic examples that are contextualised in relation to theory. The perspective of this research is influenced by postphenomenology,³ which emphasises the role played by technoscientific instruments in mediating humans’ experience of reality. Ihde importantly argues that such mediation not only mediates but also qualitatively alters perceptual experience, playing a hermeneutic role in the process. The approach of media archaeology⁴ is also influential to this research, seeking insights about current media artefacts through related historical and technological developments. A contextual understanding is especially relevant to theorising ML because it enables us to see how it is indeed novel, in addition to how it remains connected to established ideas regarding art and technology.

Mediation of the Visible

In recent years, increasing attention has been paid to the algorithmic qualities of images, what has been referred to as an “algorithmic turn”⁵ in visual media. The contrast between what is visibly apparent on the surface of images and what goes on in their subfaces is highlighted especially well in visual applications of ML in which the process involved may be highly opaque⁶ to viewers. For example, it has been proven that ML systems are capable of producing highly unpredictable, surprising results⁷ and adversarial approaches have demonstrated how images may be processed in significantly different ways by humans and machines. Adversarial approaches seek to trigger errors in ML systems. This kind of approach may be used for various purposes, including seeking to attack or compromise a system, or for diagnostic purposes, to identify and to improve potential weaknesses.

An especially notable example of the discrepancies that may arise between human and machine forms of visual processes can be found in the cross-disciplinary work of Harun Farocki.⁸ Several of Farocki’s artworks and an influential essay entitled *Phantom Images*⁹ probe the engagement of highly automated imaging systems with non-visual processes. Operative — or operational — images, Farocki says, “are images that do not represent an object, but rather are part of an operation.”¹⁰ Farocki also points out that visual technologies thus enable us to “monitor process(es) that, as a rule, cannot be observed by the human eye.”¹¹ For this reason, many have been captivated by the possibility for machine vision to act as a metaphor for an alternative to or an extension of human vision.

Trevor Paglen’s explorations with the concept of the operational image¹² often seek to visualise the invisible¹³ aspects at work in ML-produced images. This may be seen, for example in *Machine Readable Hito*,¹⁴ in which numerous portraits of the artist Hito Steyerl are displayed with labels indicating an emotion analysis of her facial expressions with a score for various categories: anger; contempt; disgust; fear; happiness; neutral; sadness; and surprise. This connects to the tradition of portraiture seeking to capture something of a sitter’s internal world through a visual representation of their face. It is also suggestive that the analysis of emotion in images by machines entails a paradox. *Training Humans*¹⁵ exhibits examples of ML training data, especially focusing on facial recognition systems. By making the image data that is typically obscured behind such systems available to viewers, the exhibition calls attention to the interplay between what is made visible or hidden away in visual processing tasks.

Rather than connecting directly to the visual and non-visual aspects of images, Hoelzl and Marie¹⁶ emphasise the “softness” of images, referring to the capacity of images to be highly variable, while adhering to strictly defined algorithmic procedures. This led, they argue, to a change from images acting as

representations of the world to taking the form of a database.¹⁷ In complement to this view, Steyerl encourages giving attention to what she refers to as the “poor image”, epitomised by networked media: “a copy in motion. Its quality is bad, its resolution substandard. As it accelerates, it deteriorates.”¹⁸ Championing data, procedure and transmissibility over the visual qualities of images is reminiscent of Farocki’s account of operative images as entailing spatial, temporal and task-based qualities. This approach enables images to be defined in ways in which go against the grain of traditional image paradigms such as painting and, to a certain extent, photography,¹⁹ which have typically championed the visual, material and humanistic qualities of images.

The works covered thus far in this paper each touch on the substantial rift that may exist between how images are produced and interpreted by machines as opposed to by humans. The participation of machines in the production – and more recently, the interpretation – of images has fuelled ongoing speculation about the potential for nonhuman forms of vision, as well as attributions of authorship to machines. Not only has it been the source of controversy questioning the authorship²⁰ and value of technically produced artefacts, such as issues of materiality, seriality and labour,²¹ but also the position of machines as interpreters of visual information. This goes beyond McLuhanian²² perspectives of media as extensions of human ability and perception, with Farocki calling attention to the capacity of machine vision (MV) to act as a “displacement of the observer’s point of view”.²³ Phantom shots, for example, are “film recordings taken from a position that a human cannot normally occupy.”²⁴ In such cases, an apparatus may act as a stand-in for the human eye may, which be used for cinematic effects, but also takes on increasingly distanced forms such as the navigation of drones or mass surveillance.

Within art contexts, the myth of the machine as artist²⁵ continues to haunt technologically engaged art. This often manifests itself in the tendency to overestimate and to fetishise machine autonomy in image production. In the case of Harold Cohen’s *AARON*,²⁶ the infamous sale of a generated portrait by the group, Obvious²⁷ and the work of artists including Mario Klingemann,²⁸ it is apparent that the participation of machines in image production is greatly overstated, as though it occurs autonomously from human intentionality and vision. The mythologisation of machines in art can also be found in a while more innocuous form in the anthropomorphising language and comparisons often applied to art involving ML, such as the use of MV as a metaphor for nonhuman vision. This frequently involves the development of adversarial strategies to evade detection from biometric surveillance, as can be seen in Adam Harvey’s *CV Dazzle*,²⁹ Zach Blas’s *Facial Weaponization Suite*³⁰ and Steyerl’s *How Not to be Seen. A Fucking Didactic Educational .MOV File*.³¹ Others treat AI and ML systems as characters that participate in the production of the work, such as in Amy Alexander’s *What the Robot Saw*³² and Memo Akten’s *Learning to See*.³³

In a step away from the anthropocentric aspects of visual media, Joanna

Zylinska's nonhuman photography³⁴ questions who or what images are *of, by, or for*,³⁵ underscoring the capacity for machines to produce images in the absence of direct human participation. Nonhuman photography also demonstrates how images may be inaccessible to humans to varying degrees – produced without human perception, agency or subjectivity playing a significant role in the process. This means that a given image may exceed its instantiation in forms tangible to humans, but it also entails the potential for highly automated imaging systems to displace the importance of the viewing subject. But despite its intentions, the idea of nonhuman photography faces the paradox of humans attempting to envision how nonhuman perception, agency and subjectivity may be materialised in image form. It nonetheless speaks to a recurring curiosity as to how technology may afford mediation between not only visual and non-visual but also between human and non-human, in such a way that it remains anthropocentric.

Mediation as a Tendency in Images

Beyond merely mediating human intentionality and the perceptual experience of both producer and viewers of images, technology also contributes to a view of technically produced images as the product of technoscientific methods. This, too, has a longer history than the use of ML in image production, having a notable effect on how photography has been theorised in comparison to painting. While visual verisimilitude had been an ideal in pictorial representation until the advent of photography, the apparent efficacy of the photographic process to faithfully capture visual likenesses of the world made it subject to scrutiny in comparison to the laborious and skilled nature of painting. Photography therefore struggled to gain legitimacy as an art form. But on the very same grounds, the presumed distancing of the photographic process from the intentionality of the photographer, photography also came to be seen as inherently truthful, scientific form of visual representation.

Technical and scientific methods offer particular ways of mediating the visible, but these do not ensure the accuracy of the images which are produced as a result. This is especially apparent in situations of error in ML systems, such as their demonstrated tendency toward inherent bias³⁶ as well as the examples made visible by adversarial approaches. But in the same way that the myth of the machine as artist continues to haunt technical forms of image production, so too does the idea of such methods imbuing images with a technoscientific perspective of the world. Many artists, as well as theorists, have criticised this kind of assumption, yet much like the paradox inherent in the concept of nonhuman photography – the inability to escape the human perspective – it appears equally difficult to take the empirical worldview out of highly technical approaches to image-making, such as the generation of visual content using ML. In this sense, the very mediating capacity which enables technical methods of visualisation to function also makes them subject to interpretation on the

level of that visualisation, but also in regards to their apprehension by viewers.

Conclusion

What is especially significant about the questions that current discourse on algorithmic methods of image production pose to us is how they contribute to a critical re-examination of the value systems that underpin theories on visual culture. The ideas and practices covered here may on the one hand more faithfully capture the nuances of current contexts than older conceptions of images as primarily visual, materially fixed, the product of a sole – human or machine – author and intended for a human audience. But they also make the image extremely hard to define by unsettling entrenched ideas concerning the ontological, communicative and mediating nature of current visual media. In this way, the application of ML to visual processing tasks does not constitute a distinct break with existing image paradigms, such as photography and painting, but builds upon these traditions, including their surrounding narratives. This underscores the wealth of not only mediating processes but also historical discourses, which may now be embedded in and behind images. While this investigation may open up more questions than it answers, it points to the fact that the mediation between human and machine perception and agency that occurs through imaging is of great relevance to not only how images operate, as well as what significance this has within visual culture.

References

- ¹ Mitchell, Melanie. *Artificial Intelligence: A Guide for Thinking Humans*. New York: Farrar, Straus and Giroux/Macmillan, 2019.
- ² Manovich, Lev. “What Is New Media?” In *The Language of New Media*, 18–61. Cambridge, London: MIT Press, 2001.
- ³ Ihde, Don. *Postphenomenology: Essays in the Postmodern Context*. Evanston: Northwestern University Press, 1993.
- ⁴ Huhtamo, Erkki, and Jussi Parikka. *Media Archaeology: Approaches, Applications, and Implications*. Berkeley, Los Angeles, London: University of California Press, 2011.
- ⁵ Ulricchio, William. “The Algorithmic Turn: Photosynth, Augmented Reality and the Changing Implications of the Image”. *Visual Studies* 26, no. 1 (March 2011): 25–35.
- ⁶ Brouwer, Joke, Lars Spuybroek, and Sjoerd van Tuinen, eds. *The War of Apparances: Transparency, Opacity, Radiance*. V2_, 2016.
- ⁷ Lehman, Joel, and et al. “The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities”, 2018.
- ⁸ Farocki, Harun. *Eye / Machine I-III, Auge/Maschine 1-III*. 2003 2001. Series of two-channel video installations, re-edited to single-channel video (colour, sound).

MoMA.

⁹ Farocki, Harun. “Phantom Images”. Edited by Saara Liinamaa. Translated by Brian Poole. *PUBLIC* 29 (2004): 12–22.

¹⁰ Farocki, p. 17.

¹¹ Farocki, p. 18.

¹² Paglen, Trevor. “Operational Images”. *E-Flux Journal* 59 (November 2014).

¹³ Paglen, Trevor. “Invisible Images (Your Pictures Are Looking at You)”. *The New Inquiry*, 8 December 2016.

¹⁴ Paglen, Trevor. *Machine Readable Hito*. 2017. Installation.

¹⁵ Crawford, Kate, and Trevor Paglen. 2019. *Training Humans*. Photography exhibition of machine learning training images.

¹⁶ Hoelzl, Ingrid, and Rémi Marie. *Softimage: Towards a New Theory of the Digital Image*. Bristol: Intellect Ltd., 2015.

¹⁷ Hoelzl and Marie, p. 96.

¹⁸ Steyerl, Hito. “In Defense of the Poor Image”. *E-Flux Journal* 10 (2009). p. 1.

¹⁹ The photographic paradigm is in a sense an exception to this, as it persists and takes on new forms in digital media, including those involving ML.

²⁰ Barthes, Roland. “The Death of the Author”. In *Image-Music-Text*, 142–48. London: Fontana Press, 1967.

²¹ Benjamin, Walter. “The Work of Art in the Age of Its Technological Reproducibility: Second Version”. In *The Work of Art in the Age of Its Technological Reproducibility and Other Writings on Media*, edited by Michael W. Jennings, Brigid Doherty, and Thomas Y. Levin, 2008, 19–55. Cambridge, London: Belknap Press, 1935.

²² McLuhan, Marshall. *Understanding Media: The Extensions of Man*. Cambridge, London: MIT Press, 1964.

²³ Virilio, Paul. “The Vision Machine”. In *The Vision Machine*, 59–77. Bloomington & Indianapolis: Indiana University Press, 1994. 5.

²⁴ Farocki, p. 13.

²⁵ Broeckmann, Andreas. ‘The Machine as Artist as Myth’. *Arts* 8, no. 1 (20 February 2019): 25.

²⁶ Cohen, Harold. C. 1973-2016. *AARON*. AI System.

²⁷ Obvious. 2018. *Portrait of Edmond De Belamy*. Painting.

²⁸ Klingemann, Mario. *The Butcher’s Son*. 2018. Digital image.

²⁹ Harvey, Adam. *CV Dazzle*. 2017 2011. Lookbook of styling tips to evade facial recognition.

³⁰ Blas, Zach. 2011. *Facial Weaponization Suite*. Series of masks.

³¹ Steyerl, Hito. 2013. *How Not to Be Seen. A Fucking Didactic Educational .MOV File*. Video.

³² Alexander, Amy. 2020. *What the Robot Saw*. Series of videos.

³³ Akten, Memo. *Learning to See*. 2018. Video.

³⁴ Zylinska, Joanna. *Nonhuman Photography*. Cambridge, London: MIT Press, 2017.

³⁵ Zylinska, p. 5.

³⁶ Buolamwini, Joy, and Timnit Gebru. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification”. In *Proceedings of*

Machine Learning Research, 81:1–15, 2018.

Author Bio

Rosemary Lee is an artist and researcher who completed her PhD at the IT-University of Copenhagen. She is currently adapting her thesis *Machine Learning and Notions of the Image* into a book. Lee's work contextualises contemporary art and technology in relation to significant historical tendencies and examples.