The Birth of Computer Vision

by James E. Dobson (2023) Minneapolis (MN), University of Minnesota Press, viii-205 pp.

Charlotte Högberg
Lund University

The ways in which we engage with computer vision are now plenty and immersed in our everyday lives. For example, we encounter computer vision practices when we are forced to tick all the boxes containing a traffic light so to visit a website, as this is turned into data to train computer vision algorithms to identify objects in pictures. We are also subjected to it when we stand in line at the airport or border controls where real-time face-recognition is deployed to identify us, and when our bodies are, or are not, identified as humans passing the crosswalk by the algorithms of a "self-driving" car. Altogether, this means that the topic of James E. Dobson's book *The Birth of Computer Vision* is very timely. If anything, the relevance of this work has only further increased since the adoption of the Artificial Intelligence (AI) Act by the European Union in March of 2024. The AI Act is one example of how a lot of both scholarly and public debate is dedicated to the opportunities, risks, and development of artificial intelligence, as well as computer vision applications specifically. For example, the question of whether real-time biometric identification of humans, by face recognition in surveillance systems, should be allowed is subject to much discussion. At the same time, Dobson shows that both the technology and the scholarly field that now go by the name of computer vision have a rich and intricate history, starting at least from the 1950s, and drawing from older statistical methods. Computer vision is moreover entangled with certain epistemological and ontological constructs, as Dobson argues: "To understand computer vision is to take on its sense of the world as a particular construct, a particular metaperspective toward reality, one that is shaped by its history" (p. 51).

One of the main goals of *The Birth of Computer Vision* is to "deconstruct the absolute division that has been drawn between accounts of human and machine vision" (pp. 3-4). Initially, the author introduces theories of ways of seeing and ideas about perception, arguing for human perception as a precondition that enables the development of machine vision. Moreover, Dobson situates the study in the fields of critical algorithm studies and the history of ideas. Such a two-fold positioning is due to the fact that the book argues that algorithms should be understood, first and foremost, as ideas. Methodologically, the book is a historical investigation, looking for the genealogies of seeing technologies through historical sources. It does this by focusing on a number of algorithms, models and technologies as case studies.

TECNOSCIENZA. Italian Journal of Science & Technology Studies
15(2) pp. 165-168 ISSN 2038-3460
DOI: 10.6092/issn.2038-3460/21234

The rationale of the case selection is based on including algorithms that have had a great impact in the infancy and formative period of the research field of computer vision and that are still important for contemporary technologies and applications.

The first chapter stresses how the visual culture of the last mid-century enabled the development of image-focused technologies and computer vision. It also discusses how the discourse of computer vision presents the technology as a "neutral extraction of information from data" (pp. 29-30) that is also superior to human vision. Machine vision assumes the world as something that can be separated into different entities, with lines demarcating specific objects and spaces. Furthermore, it builds on knowledge from the past to algorithmically shape current technologies and the culture around them. To a large extent, computer vision also takes shape in military research labs, with the United States (US) Department of Defense as a main funder, and is spurred by the military technologies of the Cold War. One first step towards developing computer vision, described by Dobson, is the making of the Perceptron, a neural network algorithm that psychologist Frank Rosenblatt applied to enable binary classification by supervised learning, as presented in Chapter 2. This algorithm was embodied in The Mark I Perceptron, a machine for image interpretation and pattern recognition. Yet, the book discusses how the binary classification soon posed limitations and after some time, Rosenblatt framed the perceptron as a research model rather than a general system able to perform visual pattern recognition.

In the following Chapters, 3 and 4, Dobson describes how research into pattern recognition and machine learning during the 1960s and 1970s moved away from computer-assisted perception of pixel-based data, and towards analysing "sensed scenes" (pp. 100-101) and high-level symbolic representations. This is the period during which the research field of computer vision was formed. According to Dobson, it was made possible by an altered ontology of the image as consisting of a collection of features. By breaking down pictures into components and parts that could be described, in edges and features, pattern recognition and object detection became what other scholars have called more "doable" problems (e.g., Fujimura 1987). Image segmentation and "blob detection" (p. 85) made it possible to distinguish and demarcate objects. Eventually, it resulted in developments such as the first system for real-time face detection, the Viola-Jones face detector framework, published in 2001. It is comprised by a set of classifiers that look for patterns in pixels and try to match them against object patterns identified as faces (p. 119). However, Dobson emphasises that these types of technologies build on normative assumptions about how an average face should look. As the author notes:

> [C]omputer vision, in its earliest moments and in the present, cannot escape from its reliance on symbolic abstractions and the biases, exclusions, and historicity that such model activity inevitably introduces. (p. 132)

Subsequently, in Chapter 4, Dobson describes the US military-funded Shakey project which aimed to produce "Shakey the robot" (p. 135). As an embodied and mobile robot, it was made to move, or rather roll, around, necessitating the ability to "sense" its environment, walls and objects of interest. In addition, by processing images from an attached camera, it constantly updated its sense of the environment. The project led to the development of the

Hough Transform that is used to detect lines. This technology is still often used as, for example, line detection algorithms are of high importance to keep cars in their lanes in assisted driving and deployed by lane departure warning systems. In the concluding chapter of the book, "Coda", more attention is devoted to the impact that these computer vision technological developments have had, and continue to have, on contemporary applications and research (p. 165). Dobson describes OpenCV, the current main toolkit for computer vision that encompasses over 2,500 algorithms, out of which one is an implementation of Rosenblatt's perceptron. The author argues that the cases genealogically analysed – the perceptron, blob detection, template pattern matching, pictorial structures and the Hough transform – have survived the test of time and are used in current technologies. This is why it is all the more important to know their history and assess how they still impact our everyday lives.

Moreover, Dobson does a great job in showing the entangled history of computer vision and military goals, and the relationships between the aims of seeing machines and surveillance. It was the US military funding which enabled much of the computer vision research, and it was to a great extent with military goals that computer vision was developed, such as to identify enemies' advances in aerial photos, to identify targets, and to ease the work for human military photo interpretations. Yet, some of the researchers, that Dobson portrays, seem less devoted to, or even against, the militarisation of the technology. Even though Shakey the robot was portrayed as harmless, it was with military operations in mind that the research behind it was funded and formed. Dobson also portrays how the military bounds led to a backlash for the research during the protests against the Vietnam War. This is argued to formally have led to universities cutting ties with military research labs, while perhaps not so much changed in terms of research practice as main actors continued to carry out research at both organisations.

Dobson also shows how the objectives and framings have changed since the beginning of what we call computer vision. When the initial goals of recognition of patterns and creating a comprehensive machine understanding of images failed to succeed, the researchers started to break it up into smaller more manageable tasks, into features, lines, and edges. In addition, when full automation proved too hard, computer vision was promoted instead as an enabling decision-support tool helping human interpreters rather than replacing them, which was deemed valid for the subsequent expert systems.

In its entirety, the work presented in *The Birth of Computer Vision* is of relevance to several scholarly conversations in current STS. By focusing on multiple actors, both humans and machines, as well as the human-machine relations, this book relates to multiple research areas in STS. For example, it is in the same vein as the work of Adrian Mackenzie (2017) who casted a light on "machine learners" by offering an archaeology of data practices, and by carrying out a sociology of programmers. Dobson's work provides a very valuable description of the cases of specific researchers and algorithms, which demonstrates and offers deep knowledge and richness in detail. However, it would in some instances have been valuable to more clearly and decisively lift the gaze and inform the reader about the significance of that particular story for later developments and for us to better understand our current technologies.

One of the most valuable aspects of the book is how it discusses computer vision as sightless seeing, that is, as a machine's direct sensing of the environment and ability to elicit knowledge from images without a sense of sight. This is also discussed in relation to ideas about the increasing absence of a graphic output in the form of representational image outputs that are visually interpretable for humans. This connects to Jussi Parikka's (2023) analysis of images as data made for machines and algorithms. Both Dobson and Parikka refer to the artist and media theorist Harun Farocki's conceptualization of operational images (e.g., Farocki 2004) and, in particular, how these operational images challenge the generally taken-for-granted ontology of images and vision. Yet, *The Birth of Computer Vision* demonstrates the highly material aspects of computer vision. For example, by the embodiment of algorithms into the Mark I Perceptron machine, the robot Shakey or the sociomaterial impact of computer vision technologies. This materiality puts into question whether it is really valid to refer to algorithms as being foremost ideas.

The Birth of Computer Vision is of high relevance for those STS scholars focusing on analysing current digital technologies. The book succeeds in its mission to show the intricate relationship between human and machine vision, and how the latter still is dependent on the former by human labelling and descriptions of images as training data for algorithms and as some algorithms working to support human vision rather than being able to replace it. As computer vision shares much history with artificial intelligence, the book offers a much-needed deepening of the historical perspectives on AI and computer vision's development. It provides a thorough empirical account to be combined with scholarship on the sociotechnical assemblages and impacts of AI, such as Kate Crawford's (2021) *Atlas of AI*. As Dobson puts great effort into describing the military influence on computer vision research, as well as the backlash that the research received due to its military ties, the book is also of interest to those researchers studying science-technology-military relationships, the impact of research funders, and political debates about universities' connections with military operations. The book's greatest contribution is, however, the insights it provides of the intense historicity that is built into our present-day highly impactful algorithms and digital applications.

References

Crawford, Kate (2021) *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*, New Haven, Yale University Press.

Farocki, Harun (2004) *Phantom Images*, in "Public", 29, pp. 12-22.

Fujimura, Joan H. (1987) *Constructing "Do-able" Problems in Cancer Research: Articulating Alignment*, in "Social Studies of Science", 17(2), pp. 257-293.

Mackenzie, Adrian (2017) *Machine learners: Archaeology of a data practice*, Cambridge (MA), MIT Press.

Parikka, Jussi (2023) *Operational images: From the visual to the invisual*, Minneapolis, University of Minnesota Press.