

Of Machines Learning to See Lemon

John Fass

Alistair McClymont

In the seventeenth-century heyday of Dutch still life painting, exemplified by painters such as Jan Davids de Heem and Pieter Claesz lemons feature heavily as opportunities for the display of matchless technique. They also fulfilled many complex symbolic functions. For the newly enriched bourgeoisie of the Dutch Golden Age, still life paintings were subtle displays of ownership and wealth derived from the Dutch colonies in Indonesia, Sri Lanka, and Taiwan and the trade in commodities extracted from these territories to Europe. This trade was facilitated by the state-supported monopoly of the Dutch East India Company, an early example of a consolidated global corporation. Over its two-hundred-year history, the Dutch East India Company transformed from a trading company into a transglobal body with many of the characteristics of an independent state. It was able to implement its strategies through force of arms, the establishment of multiple independent markets from Amsterdam to Jakarta, and the determined exploitation of conquered lands and peoples.

Still life paintings are a product of this sociopolitical system, which accrued unprecedented levels of wealth for those with access to the opportunities provided by the company, as Toby Sonneman has pointed out in her history of the lemon. The visual work performed by fruit in these images reminds the viewer that the owner can afford to purchase and consume fruit imported from far away (and, of course, that they have sufficient spare income to engage an artist to mediate this message). Lemons, like all fruit in still life paintings, conjure impressions of decay and the transient nature of human life, but they also have a particular place in communicating acidity or bitterness, as well as the interior-exterior dynamic of hidden power exerting its influence seen in the contrast between the rough outer skin of a lemon and its glistening lustrous interior. Art historian Julie Berger Hochstrasser has emphasised the importance of pictures of lemons as especially extravagant ways of displaying affluence—“the ostentation of a whole lemon peeled and sitting at the ready, just for a little squeeze of juice” (Hochstrasser, 200).

Lemons were by no means unknown in Northern Europe, having been brought back from the Middle East by Crusaders in the eleventh century, and were certainly known in antiquity as recorded by Theophrastus' *Historia Plantarum* of 300BC. Nevertheless, in the seventeenth century, only those with the financial means to purchase and consume them would know what they were and what they tasted like. Recognising a lemon by its shape, texture, and colour—for what it was and for its cultural and symbolic significance—was only possible for those who could see real lemons in the marketplace or identify them in visual representations.

In contrast to the social reading of how physical objects are recognised, in neuroscience and computer science object recognition in the human visual system is considered to be a function of complex brain processes that depend on a cascade of reflexive operations. Di Carlo observes, “we effortlessly detect and classify objects from among tens of thousands of possibilities 7 and we do so within a fraction of a second 6 despite the tremendous variation in appearance that each object produces on our eyes.”¹ The complexity of these neural operations are evident in the fact that “All visual cortical areas share a six-layered structure and the inputs and outputs to each visual area

¹ references below

share characteristic patterns of connectivity” (DiCarlo et al. 2012) ^{1,2} The task of producing a representation of sufficiently high fidelity to visually identify objects is poorly understood, and there is considerable debate about how it takes place in the brain. Computer scientists have concentrated on constructing computational models of perception, in order to produce explanations along the lines of a Turing machine, i.e. mathematical models that can simulate an infinite number of states. Neuroscientists, in contrast, have focused on the spatial distribution of the relevant brain activity and how these areas may be connected to each other, in other words operating at the cellular and molecular level of cortical circuitry.

A computer vision system recognises a lemon in a very different way, although it is rooted in the search for the neural ‘algorithm’ of human visual object recognition. Computational object recognition requires ‘training’ using relevant data. The ability of a computer to correctly identify a lemon depends on the number, quality, and accuracy of examples in the data class *lemon* the system has been exposed to. In addition, the human visual system is very good at processing interruptions to the visual field (we have no difficulty in, say, recognising a tennis racket that is resting on a chair in front of a window that looks onto a river). In the scenario above a computer would need to identify every object, perceive how they are arranged in space, assign the correct label to each one, and recognise the whole as a scene. As van Gool points out, “The same object will look different depending on the viewpoint, the illumination, or the occlusions caused by other objects in front.”³ Furthermore, the wide variation between instances of the same object means that the “recognition of an object as belonging to a particular group is a harder problem for a computer than the recognition of a specific object.”⁴ One lemon does not look exactly like another, and therefore significant computational resources must be devoted to distinguish between different views, types, or examples of lemons as an object class. As a result, much attention is given to pre-categorisation of images of objects via tagging or other taxonomic labelling methods. This requires large amounts of individual images, but perhaps more significantly a dominant logic of categorisation, to work.

As more and more detailed models are developed based on the millions of images used as training data for object recognition learning algorithms, so system complexity increases. Torralba says, “Deep learning works very well, but it’s very hard to understand why it works—what is the internal representation that the network is building.”⁵ The outcome is an opaque system, resistant to analysis, impervious to scrutiny. Often the data scientists behind this work have no idea why they obtain certain results, and have to commit resources to reverse-engineering them in order to gain a deeper understanding.⁶ This opacity of computational recognition systems means the steps involved in recognising a lemon are transformed from a set of specific associations, enculturated by human circumstance and experiential phenomena, into the output of an impenetrable probabilistic matching algorithm. Image recognition systems are of course, the result of a set of cultural assumptions about efficiency, accuracy, and performance, usually enacted through

² Felleman and Van Essen, 1991

³ 2017

⁴ van Gool, 2017

⁵ in Hardesty, 2015

⁶ Torralba, 2015

Sonneman, Toby. *Lemon: A Global History*. London: Reaktion Books, 2012. Print.

Hochstrasser, Julie Berger. *Still Life and Trade in the Dutch Golden Age*. New Haven, CT: Yale University Press, 2007. Print.

Biederman I. Recognition-by-components: a theory of human image understanding. *Psychol Rev.* 1987;94:115–147.

Potter MC. Short-term conceptual memory for pictures. *J Exp Psychol [Hum Learn]* 1976;2:509–522.

Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex.* 1991;1:1–47.

Pinto N, Cox DD, Corda B, Doukhan D, DiCarlo JJ. Why is real-world object recognition hard?: Establishing honest benchmarks and baselines for object recognition. COSYNE; Salt Lake City, UT: 2008a.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, 73(3), 415–434. <http://doi.org/10.1016/j.neuron.2012.01.010>

Hardesty, L. (2015) How computers can learn to recognise objects, <https://www.weforum.org/agenda/2015/05/how-computers-can-learn-to-recognise-objects/>, retrieved 08/18.

Einarson, B. (1976). The manuscripts of Theophrastus' *Historia plantarum*. *Classical Philology*, 71(1), 67-76.