`kittydar`: detecting edges in the world

Adrian Mackenzie

Sociology, Lancaster University a.mackenzie@lancaster.ac.uk

Take the case of `kittydar`, a small demonstration of machine learning techniques in the area of computer vision (see kittydar): '`kittydar` is short for kitty radar. `kittydar` takes an image (canvas) and tells you the locations of all the cats in the image' (Arthur 2012). This playful piece of software demonstrates the detection of edges in the mundane, not to say banal, domain of cat photos on the internet. Heather Arthur, who developed `kittydar` writes:

> Kittydar first chops the image up into many "windows" to test for the presence of a cat head. For each window, kittydar first extracts more tractable data from the image's data. Namely, it computes the Histogram of Orient Gradients descriptor of the image, using the hog-descriptor(http://github.com/harthur/hog-descriptor) library. This data describes the directions of the edges in the image (where the image changes from light to dark and vice versa) and what strength they are. This data is a vector of numbers that is then fed into a neural network(https://github.com/harthur/brain) which gives a number from 0 to 1 on how likely the histogram data represents a cat. The neural network (the JSON of which is located in this repo) has been pre-trained with thousands of photos of cat heads and their histograms, as well as thousands of non-cats. See the repo for the node training scripts (Arthur 2012).

This toy example of machine learning finds cat heads in photographs, but we can easily locate similar techniques in use in self-driving cars, border security systems, military robots, gesture recognition systems, online advertising, supply chain logistics as well as truly legion scientific applications. Their popularization is striking, and perhaps attests to an ontological gradient, a curvature that shifts and re-shapes beings.

A couple of facets attend the advent of such techniques, whose results can be seen in Figure 1:

2. `kittydar` is alluring because it imitates something people do in the visual field. In this case, the visual form of cats, massively abundant and proliferating in contemporary visual culture, are already attachment memes which `kittydar` follows in its own ways.
3. `kittydar` and its ilk can only really group beings within rigid yet synthetic limits. As the description provided by Arthur suggests, the software finds cats by chopping up the images into smaller windows. For each window, it measures a set of gradients running from light and dark in order to find

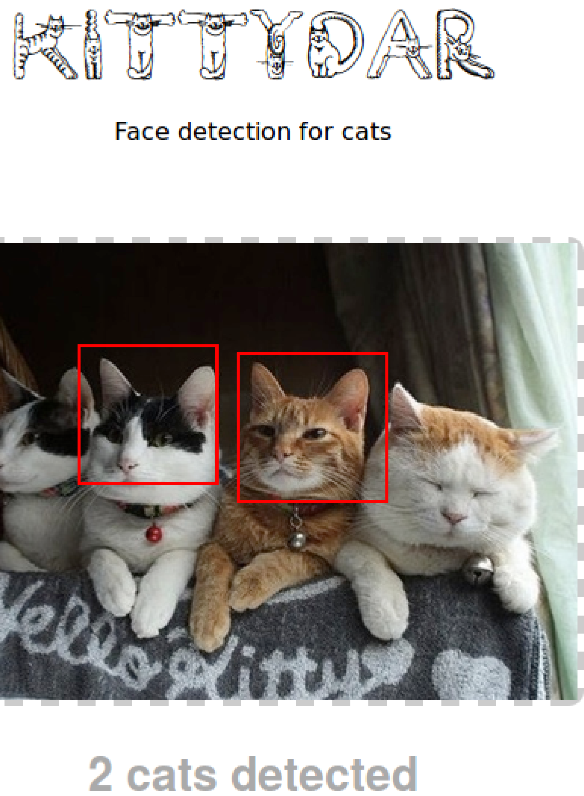

Figure 1: ‘kittydar‘ detected cat faces

'edges,' and then compares the direction of these edges to the edges found in known cat images (the training set). The intuition here is that edges are associated with the features on a cats in a fairly regular pattern.

4. The work of classification occurs somewhat generically. In this case it is either given to a neural network, a typical machine learning technique and one that has recently been heavily developed by researchers at Google (Le et al. 2011), themselves working on images of cats among other things taken from Youtube videos (BBC 2012), or to a support vector machine, a technique first developed in the 1990s by researchers working at IBM (Cortes and Vapnik 1995). Both techniques have received heavy use in science, industry, media, health and commerce, suggesting again that what we see in `kittydar` is no eccentric plaything, but a symptom of a widely generalizing classificatory process. `kittydar` runs in Javascript in a web browser. This banal execution of `kittydar` as Javascript suggests a flat and wide dispersal of techniques that were once the preserve of major research projects. The tide of cat photographs on the internet is channelled here by the surge

Imitation, classificatory rigidity, and banal generalization together configure a system of transformations that may well have some ontological significance. Each in their own way suggests forms of movement. The imitative vector follows existing patterns whose proliferation and profusion has created problems of scale. What to do with all the cat images on the internet? Build machines that locate them amidst all the other images. The classificatory rigidity and attendant errors of such machines (for instance, `kittdar` identifies some human faces as cats) could be seen as failures and inadequacies, but in fact engender renewed efforts to optimise the classificatory performance. Error is no obstacle, but like delinquency in Foucault's penal institutions (Foucault 1977) something absorbingly generative in its synthetic possibilities. Finally, these effort to optimise out errors also tends to generalize the techniques themselves. Each time a limit is addressed, each time a more complicated set of edges can be detected (cats side on for instance), the algorithm also lends itself to further generalization.

Systems of transformations, and component vectors of movement do not so much comprise a different ontology, but as diagram an ontological gradient, a surface whose variable curvature changes the things (objects, subjects, fields, institutions) moving across it. This ontological gradient brings together in a densely calculative weave widely distributed elements drawn from infrastructures, sciences, institutional practices (such as 'learning' and 'training'), and the transient and plural transients of popular culture. Amidst very powerful vectors of normalization (the identification of cat faces does not in itself suggest any radical alternatives and may well de-potentialize alternative futures), there is a slender chance that some different coherences, different configurations of beings, and perhaps some new edges might be articulated and become visible in the world.

References

Arthur, Heather. 2012. "Harthur/kittydar. GitHub." https://github.com/harthur/kittydar.

BBC. 2012. "Google 'brain' machine spots cats." *BBC News*, June. http://www.bbc.co.uk/news/technology-18595351.

Cortes, C., and V. Vapnik. 1995. "Support-Vector Networks." *Machine Learning* 20 (3): 273–97. doi:10.1023/A:1022627411411.

Foucault, Michel. 1977. *Discipline and Punish: The Birth of the Prison.* Translated by Allan Sheridan. New York: Vintage.

Le, Quoc V., Marc'Aurelio Ranzato, Rajat Monga, Matthieu Devin, Kai Chen, Greg S. Corrado, Jeff Dean, and Andrew Y. Ng. 2011. "Building high-level features using large scale unsupervised learning." *arXiv:1112.6209*, December. http://arxiv.org/abs/1112.6209.