Hallucinating canines. Google DeepDream's taxonomic heritage

↗7c38we5h ● German translation available online

Thomas Hawranke, Lasse Scherffig

Chunk 1

"Who can resist the charms of a Yorkshire Terrier? What could shake the blues from your lonely evening more readily than a blue and tan toy terrier? It would appear that almost anyone who's inclined to own a Yorkshire Terrier should do so! There are so many gigantic advantages wrapped up in this smallest of dog breeds."[1]

In 2015, dog breeds of all kinds became the face of artificial intelligence. Google released DeepDream, a technique for "peeking inside"[2] neural networks that quickly became a pop-cultural phenomenon. As such, it was primarily used to alter images into ones with a dream-like aesthetic that often would feature signs of animality: Snouts, eyes and fur were morphed into the original images, sometimes resulting in the infamous 'puppyslug' animal hybrid.

1: Dog Fancy Magazine. (2009). *Yorkshire Terrier (Smart Owner's Guide).* Freehold, NJ: Kennel Club Books, 4.

2: Mordvintsev, A., Olah, C., & Tyka, M. (2015). *Inceptionism: Going Deeper into Neural Networks.* https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html

Chunk 2

1: An image of space with puppyslugs, © Hawranke/Scherffig, 2023.

This tendency towards animality is deeply tied to the technical and epistemic foundations of the dispositive DeepDream was built in. While the actual images generated by it were a mere residue of the attempt to understand how certain neural networks operate, they became a major driving force of the popular reception of deep learning research.

A few years before, 2012 marked a turning point in the history of image recognition: With AlexNet, for the first time a Convolutional Neural Network (CNN) won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). The goal of this competition is learning classification: After looking at a vast number of images that are labelled with categories (e.g.

Chunk 3

'car' or 'dog'), a machine learns to assign them to these categories. This task poses an interesting engineering problem while at the same time raising old questions about the 'content' of images in general, because "[…] images in and of themselves have, at best, a very unstable relationship to the things [they] seem to represent, one that can be sculpted by whoever has the power to say what a particular image means."[3]

Page 88, Chunk 7: (dis-)embodied minds – creativity…
Page 49, Chunk 8: Robotik Lab (HfG Offenbach)
Page 61, Chunk 8: AI+D Lab (HfG Schwäbisch Gmünd)

Like all Neural Networks, CNNs are biologically motivated, but they are actually mathematical approximation methods. In contrast to previous neural network architectures, CNNs learn how to filter images in a way that optimally helps them to 'recognize' the categories that occur in the training data.

3: Crawford, K., & Paglen, T. (2019). *Excavating AI: The Politics of Images in Machine Learning Training Sets.* https://excavating.ai/.

4: Deng, J., Dong, W., Socher, R., Li, L., Li, K. & Fei-Fei, L. (2009). ImageNet: A Large-Scale Hierarchical Image Database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition: CVPR 2009* ; Miami [Beach], Florida, USA, 20 - 25 June 2009, 248-55. Piscataway, NJ: IEEE.

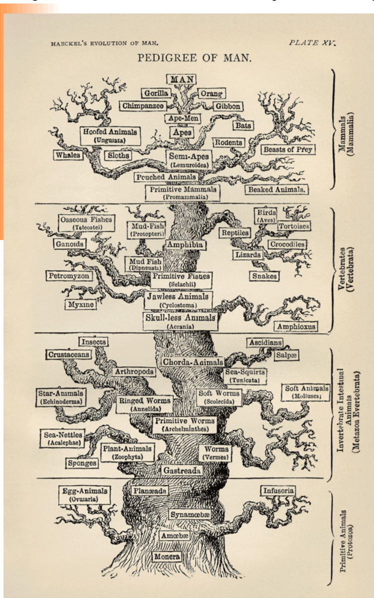
5: Crawford/Paglen, *Excavating AI*

The training data for the ILSVRC are based on images from the ImageNet library, which consists of 12 million images in 21,000 categories.[4] These images were scraped from the Internet and hand-labelled via Amazon's Mechanical Turk (MTurk). During its creation, ImageNet was the biggest academic employer on MTurk with about 25,000 workers.[5]

The categories of ImageNet, in turn, were taken from WordNet – a database of the English language that organizes all words into groups of "cognitive synonyms" called "synsets" and these into a tree-like order (e.g. from Yorkshire Terrier to mammal to animal to organism).

For ImageNet, the use of synsets "takes care of disambiguating word meanings and of combining together synonyms into the same object category."[6] Chunk 4

6: Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C. & Fei-Fei, Li. (2015). ImageNet Large Scale Visual Recognition Challenge. *Int J Comput Vis 115,* (3), 211–52. https://doi.org/10.1007/s11263-015-0816-y.



2: Pedigree of man (Haeckel 1874), © public domain.

The ILSVRC Chunk 5 hence is firmly situated within a history of Western attempts to organize knowledge into taxonomies – attempts that originally would place 'man' at the top of a tree of life and thus reveal their ideological nature (fig. 2).

This history again shows, as Caillois has noted, the unstable relationship between things and their categorization: Chunk 6

"Before classifying vertebrates as mammals, birds, batrachians, reptiles, or fish, they were grouped according to the number of feet they had. Horses were put into the same category as frogs and turtles. [...] Nonetheless, it should be said that having four feet is an interesting feature as well, with certain specific and ineluctable consequences, which is almost eliminated as an object of study, though, by the new, improved taxonomy.

Residual characteristics that have been legitimately disqualified surely give rise to remarkable relationships that are indubitably worth detecting and establishing. Chunk 7

Even though they have been excluded, they are by no means insignificant. Chunk 8

Chunk 9

7: Caillois, R. (2003). *The Edge of Surrealism: A Roger Caillois Reader* /by Roger Caillois. Ed. And with an Introduction by Claudine Frank. Transl. By Claudine Frank and Camille Naish. Durham: Duke Univ. Press. 344-345.

8: Russakovsky et al., *ImageNet Large Scale Visual Recognition Challenge*.

9: Russakovsky et al. (2015). For the top ten categories within the dataset see: Thoma, Martin. 2016. *What Is the Distribution of Categories in Imagenet Training Set (ILSVRC2012)*. https://datascience.stackexchange.com/questions/11777/what-is-the-distribution-of-categories-in-imagenet-training-set-ilsvrc2012.

10: Connor, M. (2015). *Why Is Deep Dream Turning the World into a Doggy Monster Hellscape?* https://rhizome.org/editorial/2015/jul/10/deep-dream-doggy-monster/

11: Connor, *Why Is Deep Dream Turning the World into a Doggy Monster Hellscape?*

[...] The universe is radiant. It supports any secant, median, chord, or bisectrix."[7]

The ILSVRC was created in 2010 as an annual benchmark for image classification, but its categories significantly changed in 2012: "In the first year of the challenge synsets were selected randomly [...]. In ILSVRC2012, 90 synsets were replaced with categories corresponding to dog breeds to allow for evaluation of more fine-grained object classification."[8] The 2012 challenge consisted of 1.2 million images belonging to 1000 object categories. Among those categories, 120 were dog breeds. The ten categories with the most images in the dataset included five of dogs, with the Yorkshire Terrier at the very top.[9] In a way, ILSVRC became a dog-detection contest.

This is probably no accident. On the one hand, dogs are cute and the designers of the ILSVRC may have chosen to inject some cuteness into their very technical challenge.[10] On the other hand, dog breeds themselves are the result not only of thousands of years of artificial selection by humans, but also of standardization. The Kennel Club, praising the "charms of a Yorkshire Terrier" above, has been one of the first international registries of dog breeds that ensure these charms will consistently be a feature of all Yorkshire Terriers and their offspring. In other words, dogs are not only a more "happy medium"[11] to train machine learning models on.

Chunk 10

Through their standardization they also make the fine-grained classification possible that otherwise would face the inherent hybridity of Caillois' radiant universe.

Chunk 11

After AlexNet, CNNs became the standard recipe for winning at ILSVRC. These systems essentially perform a series of calculations to convert a two-dimensional image into a single vector encoding the probability to which of the 1000 categories it belongs.

Chunk 12

12: Mordvintsev et al., *Inceptionism: Going Deeper into Neural Networks*

13: This corresponds with the idea of AI art as signal processing as put forward by Fabian Offert in this book.

In consequence, researchers became interested in "what exactly goes on at each layer" of this hierarchical pipeline.[12]

DeepDream was one attempt at understanding this. The technique consisted of optimizing the input to a CNN in order to maximize the activation of the network, or a part of it. The method became most popular when used to analyse an image and then maximize the features found in it – thereby 'hallucinating' whatever the network tended to detect in the image.

DeepDream, originally conceived to analyse Google's Inception CNN, may have started current generative AI but its generative potential was a side effect of the attempt to understand CNNs. The images it produced were images whose 'content' was the functioning of the network itself. [13]

Chunk 13

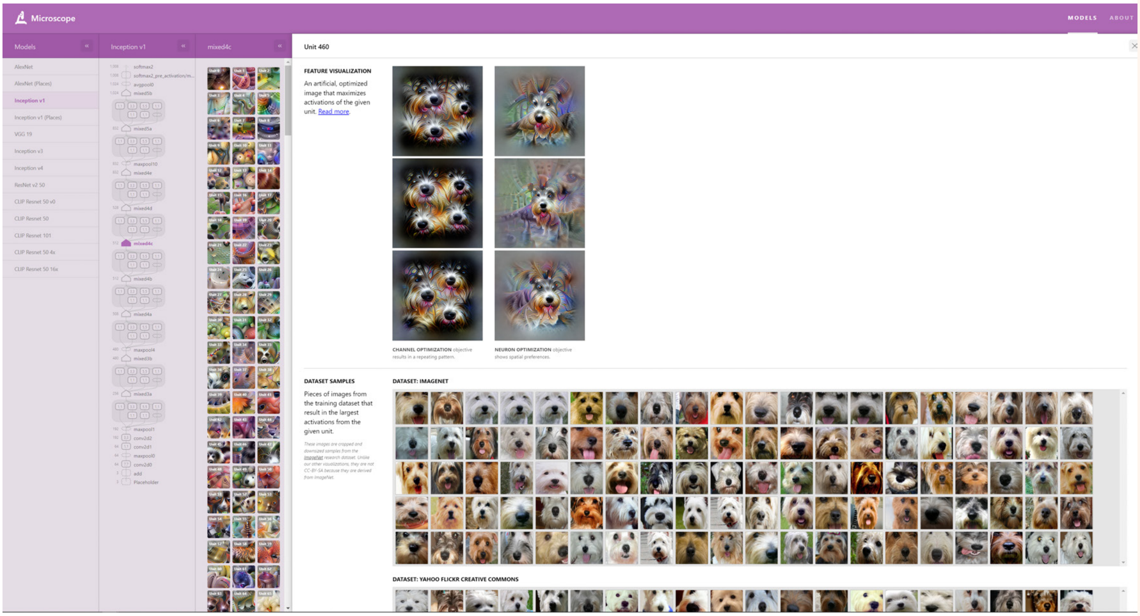
Because of the fine-grained detection challenge added in 2012, this content became heavily dog-focused.

Chunk 14

As CNNs learn to filter images for classification, the filters they learn reflect those features in the training-data which are useful for discrimination. Because they had to be able to detect the difference between a Yorkshire Terrier and other dog breeds and because the training data proportionally featured more dogs than other 'objects,' the networks learned the invariant features of dog photography.

They did not learn to detect dog breeds, but became embodiments of how we depict them – even forming specific ‘dog neurons’ in the process (fig. 3). When DeepDream turned the classification process upside-down by maximizing the features found in an image, it consequently became an image generator with a dog bias.

Chunk 15



3: A dog neuron in the OpenAI Microscope, © CC-BY 4.0, 2023, https://microscope-azure-edge.openai.com/models/inceptionv1/mixed4c_0/460

It is only fitting that the images ‘hallucinated’ by DeepDream not only reflect the composition of the ImageNet dataset, but also its situatedness in a history of classification and standardization. DeepDream hence reinforced the standardization of dog cuteness which, in turn, became the face of progress in machine learning while illustrating its ideological roots. Describing the historical development of taxonomy in eighteenth- and nineteenth-century Britain, Harriet Ritvo points out the drive to appropriate foreign species into the domestic order. Like every form of appropriation, this not only runs one way: “the classification of animals, like that of any group of significant objects, is apt to tell as much about the classifiers as about the classified.”[14]

14: Ritvo, H. (1998). *The Platypus and the Mermaid, and Other Figments of the Classifying Imagination*. 1st Harvard University Press pbk. ed. Cambridge, Mass. Harvard University Press., xii.

This article is separately published as DOI: 10.25358/openscience-10777.