thesis

Depths of learning

The coming revolution in smart machines and robots has been predicted by artificial intelligence enthusiasts in every decade since the 1950s. It's never happened. Progress in artificial intelligence has been halting, tedious, in many ways disappointing. Notable successes — IBM's Deep Blue computer beating Gary Kasparov in chess in 1997, and the emergence of reasonably useful language translation in the late 1990s mostly came through brute computation, rather than any essential similarity between computer operations and the mechanisms behind human thought or perception.

But this history of failure is now changing, and sharply. A decade ago, algorithms for computer vision and object recognition could barely identify a simple ball or block on a plain background. Now they can distinguish distinct human faces more or less as well as real people, even in complex natural backgrounds. Two months ago, Google released a smartphone app capable of translating text from more than 20 foreign languages just from photos — of a road sign, perhaps, or a handwritten note. One of Google's research groups recently showed that an algorithm — simply by watching and learning - could play all of the Atari video games of the 1980s with the same skill as human experts.

It's all the result of the discovery that some old ideas about neural networks, if altered in small but significant ways, are much more powerful than anyone expected — and of a conscious effort to mimic some of the good tricks used in the biology of human and animal perception. This time the revolution in artificial intelligence looks to be for real.

Machine-learning research has always been about getting machines to learn to recognize complex patterns. Connect a computer to a camera, and a good image identification algorithm should be able, from the raw image data, to identify a particular human face, a coffee cup, or a Husky pulling a sled. Historically, however, such networks have always struggled. Even rudimentary success demanded human intervention people helping the algorithms by defining important large-scale features that the algorithms could look for, such as edges or simple geometrical figures. Algorithms couldn't learn to see the usefulness of these features on their own.

Not any more. This has all changed owing to the discovery of so-called deep learning neural networks, which act on raw image



This time the revolution in artificial intelligence looks to be for real.

data, in much the same way as the human eve and brain. These neural networks take the raw pixel data as input to a first layer of neuronal elements, and process it through nonlinear connections to multiple further layers of other neurons. After training, these come to represent more abstract aspects of the image. In the case of vision, for example, the next layer may capture features such as edges and their orientations in particular locations. A further layer might then detect motifs built from edges, and another might assemble such motifs into larger patterns reflecting parts of familiar objects. Impressively, such networks can learn to recognize the most useful features of images without any human help.

Writing in Nature, Yann LeCun, Yoshua Bengio and Geoffrey Hinton offer an excellent review of recent progress (see Nature 521, 436-444; 2015). A crucial problem that these methods overcome, they note, is linked to the basic symmetries of objects in images. Images of a Husky in two different poses ought to register as the same dog. Yet, in a pixel-by-pixel description, such images might be almost completely different. Likewise, images of a Husky and Samoyed in the same pose, which should be distinguished, might be almost identical. The 'deep' in deep learning refers to the number of input-to-output layers in the network, which now often reaches 10 or 20, rather than only a handful in earlier work. As researchers have discovered, with multiple nonlinear layers, such systems can express intricate functions of their inputs and so remain sensitive to small feature details while being insensitive to image variations linked to basic symmetries — changes in lighting, pose or background, for example.

Perhaps the most exciting advance of all involves networks that mimic aspects of human and animal perception. In the visual cortex of the human brain, cells respond to light striking small, overlapping subregions of the visual field. These neurons then feed into further higher processing layers in a hierarchical way. Neural networks following this design — so-called 'convolutional' neural networks — have inputs representing small, overlapping regions of an image or text, and are, as a result, well tuned to structural hierarchies present in natural signals. In images, combinations of edges form motifs, which then assemble into parts, whereas in speech or text, basic sounds integrate to phonemes, then syllables, words and sentences.

As LeCun and colleagues point out, these neural networks, because of their architecture, actually respond to images in the way real brains do. Studies find, for example, that the neuron-by-neuron activity pattern in a neural net trained on an image accounts for about half the variance found in random sets of neurons in the visual cortex of a monkey viewing the same image. This structural trick is now used in most applications of machine learning for recognition and detection not only for text and images, but in areas such as identifying potentially useful drug molecules or analysing particle accelerator data. Advancing technology has helped, as modern architectures work with as many as 10 to 20 layers of nonlinear elements, training through the adjustment of hundreds of millions of weight parameters.

These breakthroughs in computer science and applied mathematics now drive frenetic research by commercial firms such as Google and Facebook. Artificial intelligence isn't what it used to be. Gill Pratt, program director for robotics research at the Defense Advanced Research Projects Agency (DARPA), suggests that artificial intelligence may soon experience an explosive event not unlike the Cambrian explosion in evolutionary history when, in a short span of time, evolution proliferated from simple, rudimentary single-celled organisms into myriad multicellular forms (J. Econ. Perspect. 29, 51-60; 2015). That explosion happened when evolution somehow stumbled on a design pathway leading out of the single-celled trap and into new open space.

In artificial intelligence, the combination of deep learning methods, advancing neuroscience and computing power is doing the same for brains and intelligence, which are no longer limited to the biological kind. There may be many further, and higher, levels of intelligence just waiting to be discovered.

MARK BUCHANAN