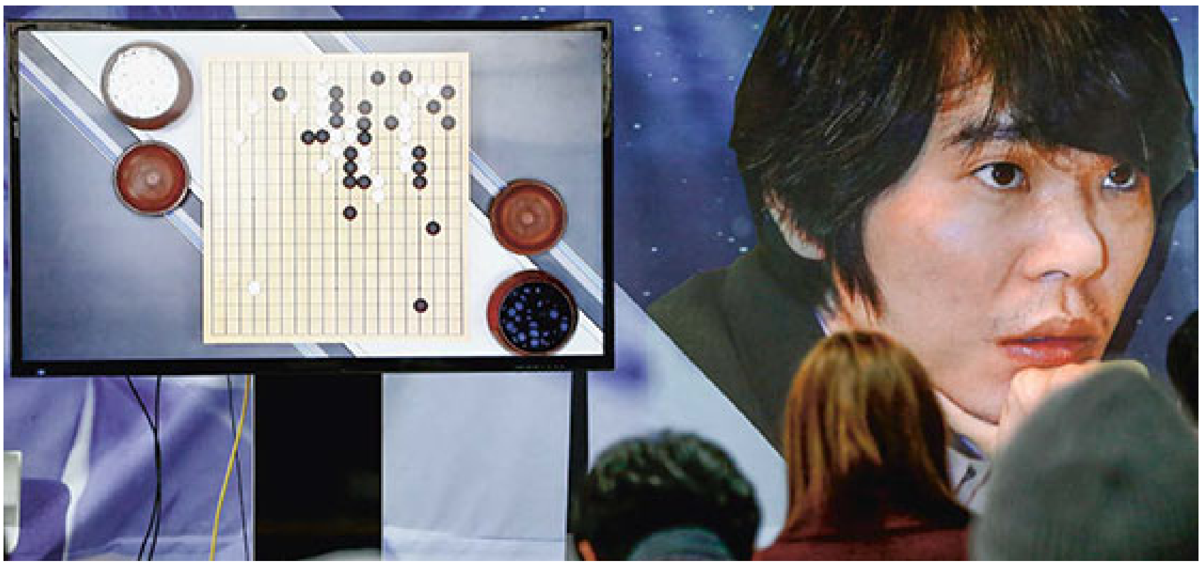
Imitation neurones, genuine potential

June 27 2016, by Fabien Goubet



Credit: Swiss National Science Foundation

In March 2016, the world Go champion Lee Sedol lost 1-4 against the artificial intelligence AlphaGo. For many, this was yet another defeat for humanity at the hands of the machines. Indeed, the success of the AlphaGo software was forged in an area of artificial intelligence that has seen huge progress over the last decade. Deep learning, as it's called, uses artificial neural networks to process algorithmic calculations. This software architecture therefore mimics biological neural networks.

Much of the progress in deep learning is thanks to the work of Jürgen Schmidhuber, director of the IDSIA (Istituto Dalle Molle di Studi sull'Intelligenza Artificiale) which is located in the suburbs of Lugano.

The IDSIA doctoral student Shane Legg and a group of former colleagues went on to found DeepMind, the startup acquired by Google in early 2014 for USD 500 million. The DeepMind algorithms eventually wound up in AlphaGo.

"Schmidhuber is one of the best at deep learning," says Boi Faltings of the EPFL Artificial Intelligence Lab. "He never let go of the need to keep working at it." According to Stéphane Marchand-Maillet of the University of Geneva computing department, "he's been in the race since the very beginning."

Cat photos, thousands of cat photos

The real strength of deep learning is structural recognition, and winning at Go is just an illustration of this, albeit a rather resounding one. Elsewhere, and for some years now, we have seen it applied to an entire spectrum of areas, such as visual and vocal recognition, online translation tools and smartphone personal assistants. One underlying principle of machine learning is that algorithms must first be trained using copious examples. Naturally, this has been helped by the deluge of user-generated content spawned by smartphones and web 2.0, stretching from Facebook photo comments to official translations published on the Internet. By feeding a machine thousands of accurately tagged images of cats, for example, it learns first to recognise those cats and later any image of a cat, including those it hasn't been fed.

Deep learning isn't new; it just needed modern computers to come of age. As far back as the early 1950s, biologists tried to lay out formal principles to explain the working of the brain's cells. In 1956, the psychologist Frank Rosenblatt of the New York State Aeronautical Laboratory published a numerical model based on these concepts, thereby creating the very first artificial neural network. Once integrated into a calculator, it learned to recognise rudimentary images.

"This network only contained eight neurones organised in a single layer. It could only recognise simple characters," says Claude Touzet of the Adaptive and Integrative Neuroscience Laboratory of Aix-Marseille University. "It wasn't until 1985 that we saw the second generation of artificial neural networks featuring multiple layers and much greater performance." This breakthrough was made simultaneously by three researchers: Yann LeCun in Paris, Geoffrey Hinton in Toronto and Terrence Sejnowski in Baltimore.

Byte-size learning

In multilayer networks, each layer learns to recognise the precise visual characteristics of a shape. The deeper the layer, the more abstract the characteristics. With cat photos, the first layer analyses pixel colour, and the following layer recognises the general form of the cat. This structural design can support calculations being made upon thousands of layers, and it was this aspect of the architecture that gave rise to the name 'deep learning'.

Marchand-Maillet explains: "Each artificial neurone is assigned an input value, which it computes using a mathematical function, only firing if the output exceeds a pre-defined threshold." In this way, it reproduces the behaviour of real neurones, which only fire and transmit information when the input signal (the potential difference across the entire neural circuit) reaches a certain level. In the artificial model, the results of a single layer are weighted, added up and then sent as the input signal to the following layer, which processes that input using different functions, and so on and so forth.

For example, if a system is trained with great quantities of photos of apples and watermelons, it will progressively learn to distinguish them on the basis of diameter, says Marchand-Maillet. If it cannot decide (e.g., when processing a picture of a tiny watermelon), the subsequent layers take over by analysing the colours or textures of the fruit in the photo, and so on. In this way, every step in the process further refines the assessment.

Video games to the rescue

For decades, the frontier of computing held back more complex applications, even at the cutting edge. Industry walked away, and deep learning only survived thanks to the video games sector, which eventually began producing graphics chips, or GPUs, with an unprecedented power at accessible prices: up to 6 teraflops (i.e., 6 trillion calculations per second) for a few hundred dollars. "There's no doubt that it was this calculating power that laid the ground for the quantum leap in deep learning," says Touzet. GPUs are also very good at parallel calculations, a useful function for executing the innumerable simultaneous operations required by neural networks.

Although image analysis is getting great results, things are more complicated for sequential data objects such as natural spoken language and video footage. This has formed part of Schmidhuber's work since 1989, and his response has been to develop recurrent neural networks in which neurones communicate with each other in loops, feeding processed data back into the initial layers.

Such sequential data analysis is highly dependent on context and precursory data. In Lugano, networks have been instructed to memorise the order of a chain of events. Long Short Term Memory (LSTM) networks can distinguish 'boat' from 'float' by recalling the sound that preceded 'oat' (i.e., either 'b' or 'fl'). "Recurrent neural networks are more powerful than other approaches such as the Hidden Markov models," says Schmidhuber, who also notes that Google Voice integrated LSTMs in 2015. "With looped networks, the number of layers is potentially infinite," says Faltings.

For Schmidhuber, deep learning is just one aspect of artificial intelligence; the real thing will lead to "the most important change in the history of our civilisation." But Marchand-Maillet sees deep learning as "a bit of hype, leading us to believe that artificial intelligence can learn anything provided there's data. But it's still an open question as to whether deep learning can really be applied to every last domain."

Provided by Swiss National Science Foundation