

## Novel math could bring machine learning to the next level

September 2 2019





The new approach allows artificial intelligence to learn to recognize transformed images much faster. Credit: Diogo Matias

A team of Italian mathematicians, including a neuroscientist from the Champalimaud Centre for the Unknown (CCU), in Lisbon, Portugal, has shown that artificial vision machines can learn to recognize complex images more quickly by using a mathematical theory that was developed 25 years ago by one of this new study's co-authors. Their results have been published in the journal *Nature Machine Intelligence*.

In recent decades, machine vision performance has vastly improved. Artificial systems can now learn to recognize virtually any <u>human face</u> or to identify any individual fish moving in a tank.

Such <u>machines</u> are, in fact, electronic models of networks of biological neurons, and their aim is to simulate the functioning of the brain, which excels at these visual tasks without any conscious effort on our part.

But how do these artificial neural networks actually learn? In the case of face recognition, for instance, they do it by acquiring experience about what human faces look like in the form of a series of portraits. More specifically, after being digitized into a matrix of pixel values, each image is "crunched" inside the <u>neural network</u>, which then extracts general, meaningful features from the set of sample faces (such as the eyes, mouth, nose, etc).

This <u>deep learning</u> enables the machine to spit out another set of values, which will in turn enable it to identify a face it has never seen before in a databank of faces (much like a fingerprint database), and therefore predict who that face belongs to with great accuracy.



## The story of Clever Hans

But before the neural network can perform this well, it is typically necessary to present it with thousands of faces (i.e. matrices of numbers). Moreover, though these machines have been increasingly successful at pattern recognition, the fact is that nobody really knows what goes on inside them as they learn tasks. They are basically black boxes.

What this means is that it is not possible to determine which or how many features the machine is actually extracting from the initial data—and not even how many of those features are really meaningful for face recognition.

"To illustrate this, consider the paradigm of the wise horse," says first author of the study Mattia Bergomi, who works in the Systems Neuroscience Lab at the CCU. The story, from the early years of the 20th century, concerns a horse in Germany called Clever Hans that his master claimed had learned to perform arithmetic and announce the result of additions, subtractions, etc. by stamping one of his front hooves on the ground the right number of times. Many people were convinced he could count; the horse was even reported by the *New York Times*. But then, in 1907, a German psychologist showed that the horse was, in fact, picking up unconscious cues in his master's body language that were telling it when to stop tapping.

"It's the same with machine learning; there is no control over how it works, or what it has learned during training," Bergomi explains. The machine, having no a priori knowledge of faces, just somehow does its stuff—and it works.

This led the researchers to ask whether there might be a way to inject some knowledge of the real world about faces or other objects into the



neural network before training in order to cause it explore a more limited space of possible features instead of considering them all—including those that are impossible in the real world. "We wanted to control the space of learned features," Bergomi says. "It's similar to the difference between a mediocre chess player and an expert: The first sees all possible moves, while the latter only sees the good ones," he adds.

Another way of putting it, he says, is by saying that "our study addresses the following simple question: When we train a deep neural network to distinguish road signs, how can we tell the network that its job will be much easier if it only has to care about simple geometrical shapes such as circles and triangles?"

The scientists reasoned that this approach would substantially reduce training time—and importantly, give them a hint about what the machine might be doing to obtain its results. "Allowing humans to drive the learning process of learning machines is fundamental to move toward a more intelligible artificial intelligence and reduce the skyrocketing cost in time and resources that current neural networks require in order to be trained," he says.

## What's in a shape?

An abstract <u>mathematical theory</u> called topological data analysis (TDA) was key. The first steps in the development of TDA were taken in 1992 by the Italian mathematician Patrizio Frosini, co-author of the new study, currently at the University of Bologna. "Topology is one of the purest forms of math," says Bergomi. "And until recently, people thought that topology would not be applicable to anything concrete for a long time, until TDA became well-known in the last few years."

Topology is a sort of extended geometry that, instead of measuring lines and angles in rigid shapes (such as triangles, squares, cones, etc.), seeks



to classify highly complex objects according to their shape. For a topologist, for example, a donut and a mug are the same object: one can be deformed into the other by stretching or compression.

Now, the thing is, current neural networks are not good at topology. For instance, they do not recognize rotated objects. To them, the same object will look completely different every time it is rotated. That is precisely why the only solution is to make these networks "memorize" each configuration separately—by the thousands. And it is precisely what the authors were planning to avoid by using TDA.

Think of TDA as being a mathematical tool for finding meaningful internal structure (topological features), in any complex object that can be represented as a huge set of numbers. This is accomplished by looking at the data through certain well-chosen "lenses," or filters. The data itself can be about faces, financial transactions or cancer survival rates. TDA makes it possible to teach a neural network to recognize faces without having to present it with each of the different orientations faces might assume in space. The machine will now recognize all faces as being a face, even in different rotated positions.

In their study, the scientists tested the benefits of combining machine learning and TDA by teaching a neural network to recognize handwritten digits. The results speak for themselves.

As these networks are bad topologists and handwriting can be very ambiguous, two different hand-written digits may prove indistinguishable for current machines—and conversely, they may identify two instances of the same hand-written digit as different. The task requires presenting the network, which knows nothing about digits in the real world, with thousands of images of each of the 10 digits written with all sorts of slants, calligraphies, etc.



To inject knowledge about digits, the team built a set of a priori features that they considered meaningful—in other words, a set of "lenses" through which the network would see the digits—and forced the machine to choose among these lenses to look at the images. The number of images (that is, the time) needed for the TDA-enhanced neural network to learn to distinguish fives from sevens, however badly written, while maintaining its predictive power, dropped down to less than 50.

"What we mathematically describe in our study is how to enforce certain symmetries, and this provides a strategy to build machine-learning agents that are able to learn salient features from a few examples by taking advantage of the knowledge injected as constraints," says Bergomi.

Does this mean that the inner workings of learning machines that mimic the brain will become more transparent in the future, enabling new insights on the inner workings of the brain itself? In any case, this is one of Bergomi's goals. "The intelligibility of artificial intelligence is necessary for its interaction and integration with biological intelligence," he says. He is currently working, in collaboration with his colleague Pietro Vertechi, on developing a new kind of neural <u>network</u> architecture that will allow humans to swiftly inject high-level knowledge into these networks to control and speed up their training.

**More information:** Towards a topological–geometrical theory of group equivariant non-expansive operators for data analysis and machine learning, *Nature Machine Intelligence* (2019). DOI: 10.1038/s42256-019-0087-3

Provided by Champalimaud Centre for the Unknown

Citation: Novel math could bring machine learning to the next level (2019, September 2)



retrieved 1 May 2024 from https://techxplore.com/news/2019-09-math-machine.html

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.