

Recent advances give theoretical insight into why deep learning networks are successful

July 28 2020, by Sabbi Lall



Credit: CC0 Public Domain

Deep learning systems are revolutionizing technology around us, from voice recognition that pairs you with your phone to autonomous vehicles that are increasingly able to see and recognize obstacles ahead. But much of this success involves trial and error when it comes to the deep learning networks themselves. A group of MIT researchers recently reviewed their contributions to a better theoretical understanding of deep learning

networks, providing direction for the field moving forward.

"Deep learning was in some ways an accidental discovery," explains Tommy Poggio, investigator at the McGovern Institute for Brain Research, director of the Center for Brains, Minds, and Machines (CBMM), and the Eugene McDermott Professor in Brain and Cognitive Sciences. "We still do not understand why it works. A [theoretical framework](#) is taking form, and I believe that we are now close to a satisfactory theory. It is time to stand back and review recent insights."

Climbing data mountains

Our current era is marked by a superabundance of data—data from inexpensive sensors of all types, text, the internet, and large amounts of genomic data being generated in the life sciences. Computers nowadays ingest these multidimensional datasets, creating a set of problems dubbed the "curse of dimensionality" by the late mathematician Richard Bellman.

One of these problems is that representing a smooth, high-dimensional function requires an astronomically large number of parameters. We know that [deep neural networks](#) are particularly good at learning how to represent, or approximate, such complex data, but why? Understanding why could potentially help advance deep learning applications.

"Deep learning is like electricity after Volta discovered the battery, but before Maxwell," explains Poggio, who is the founding scientific advisor of The Core, MIT Quest for Intelligence, and an investigator in the Computer Science and Artificial Intelligence Laboratory (CSAIL) at MIT. "Useful applications were certainly possible after Volta, but it was Maxwell's theory of electromagnetism, this deeper understanding that then opened the way to the radio, the TV, the radar, the transistor, the computers, and the internet."

The theoretical treatment by Poggio, Andrzej Banburski, and Qianli Liao points to why deep learning might overcome data problems such as "the curse of dimensionality." Their approach starts with the observation that many natural structures are hierarchical. To model the growth and development of a tree doesn't require that we specify the location of every twig. Instead, a model can use local rules to drive branching hierarchically. The primate visual system appears to do something similar when processing complex data. When we look at natural images—including trees, cats, and faces—the brain successively integrates local image patches, then small collections of patches, and then collections of collections of patches.

"The physical world is compositional—in other words, composed of many local physical interactions," explains Qianli Liao, an author of the study, and a graduate student in the Department of Electrical Engineering and Computer Science and a member of the CBMM. "This goes beyond images. Language and our thoughts are compositional, and even our nervous system is compositional in terms of how neurons connect with each other. Our review explains theoretically why deep networks are so good at representing this complexity."

The intuition is that a hierarchical neural [network](#) should be better at approximating a compositional function than a single "layer" of neurons, even if the total number of neurons is the same. The technical part of their work identifies what "better at approximating" means and proves that the intuition is correct.

Generalization puzzle

There is a second puzzle about what is sometimes called the unreasonable effectiveness of deep networks. Deep network models often have far more parameters than data to fit them, despite the mountains of data we produce these days. This situation ought to lead to

what is called "overfitting," where your current data fit the model well, but any new data fit the model terribly. This is dubbed poor generalization in conventional models. The conventional solution is to constrain some aspect of the fitting procedure. However, deep networks do not seem to require this constraint. Poggio and his colleagues prove that, in many cases, the process of training a deep network implicitly "regularizes" the solution, providing constraints.

The work has a number of implications going forward. Though deep learning is actively being applied in the world, this has so far occurred without a comprehensive underlying theory. A theory of deep learning that explains why and how deep networks work, and what their limitations are, will likely allow development of even much more powerful learning approaches.

"In the long term, the ability to develop and build better intelligent machines will be essential to any technology-based economy," explains Poggio. "After all, even in its current—still highly imperfect—state, [deep learning](#) is impacting, or about to impact, just about every aspect of our society and life."

More information: Tomaso Poggio et al. Theoretical issues in deep networks, *Proceedings of the National Academy of Sciences* (2020). [DOI: 10.1073/pnas.1907369117](https://doi.org/10.1073/pnas.1907369117)

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: Recent advances give theoretical insight into why deep learning networks are successful

(2020, July 28) retrieved 4 May 2024 from <https://techxplore.com/news/2020-07-advances-theoretical-insight-deep-networks.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.