Framework improves 'continual learning' for artificial intelligence

May 15 2019, by Matt Shipman

Credit: SilverBlu3

Researchers have developed a new framework for deep neural networks that allows artificial intelligence (AI) systems to better learn new tasks while "forgetting" less of what it has learned regarding previous tasks. The researchers have also demonstrated that using the framework to learn a new task can make the AI better at performing previous tasks, a phenomenon called backward transfer.

"People are capable of continual learning; we learn new tasks all the time, without forgetting what we already know," says Tianfu Wu, an assistant professor of electrical and computer engineering at NC State and co-author of a paper on the work. "To date, AI systems using deep neural networks have not been very good at this."

"Deep neural network AI systems are designed for learning narrow tasks," says Xilai Li, a co-lead author of the paper and a Ph.D. candidate at NC State. "As a result, one of several things can happen when learning new tasks. Systems can forget old tasks when learning new ones, which is called catastrophic forgetting. Systems can forget some of the things they knew about old tasks, while not learning to do new ones as well. Or systems can fix old tasks in place while adding new tasks – which limits improvement and quickly leads to an AI system that is too large to operate efficiently. Continual learning, also called lifelong-learning or learning-to-learn, is trying to address the issue."

"We have proposed a new framework for continual learning, which decouples network structure learning and model parameter learning," says Yingbo Zhou, co-lead author of the paper and a research scientist at Salesforce Research. "We call it the Learn to Grow framework. In experimental testing, we've found that it outperforms previous approaches to continual learning."

To understand the Learn to Grow framework, think of deep neural networks as a pipe filled with multiple layers. Raw data goes into the top of the pipe, and task outputs come out the bottom. Every "layer" in the pipe is a computation that manipulates the data in order to help the network accomplish its task, such as identifying objects in a digital image. There are multiple ways of arranging the layers in the pipe, which correspond to different "architectures" of the network.

When asking a deep neural network to learn a new task, the Learn to

Grow framework begins by conducting something called an explicit neural architecture optimization via search. What this means is that as the network comes to each layer in its system, it can decide to do one of four things: skip the layer; use the layer in the same way that previous tasks used it; attach a lightweight adapter to the layer, which modifies it slightly; or create an entirely new layer.

This architecture optimization effectively lays out the best topology, or series of layers, needed to accomplish the new task. Once this is complete, the network uses the new topology to train itself on how to accomplish the task – just like any other deep learning AI system.

"We've run experiments using several datasets, and what we've found is that the more similar a new task is to previous tasks, the more overlap there is in terms of the existing layers that are kept to perform the new task," Li says. "What is more interesting is that, with the optimized – or "learned" topology – a network trained to perform new tasks forgets very little of what it needed to perform the older tasks, even if the older tasks were not similar."

The researchers also ran experiments comparing the Learn to Grow framework's ability to learn new tasks to several other continual learning methods, and found that the Learn to Grow framework had better accuracy when completing new tasks.

To test how much each network may have forgotten when learning the new task, the researchers then tested each system's accuracy at performing the older tasks – and the Learn to Grow framework again outperformed the other networks.

"In some cases, the Learn to Grow framework actually got better at performing the old tasks," says Caiming Xiong, the research director of Salesforce Research and a co-author of the work. "This is called

backward transfer, and occurs when you find that learning a new task makes you better at an old task. We see this in people all the time; not so much with AI."

The paper, "Learn to Grow: A Continual Structure Learning Framework for Overcoming Catastrophic Forgetting," will be presented at the 36th International Conference on Machine Learning, being held June 9-15 in Long Beach, California.

More information: Learn to Grow: A Continual Structure Learning Framework for Overcoming Catastrophic Forgetting. arxiv.org/pdf/1904.00310.pdf

Provided by North Carolina State University