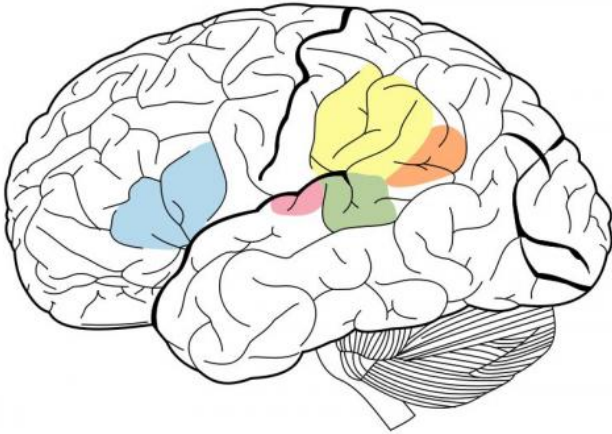


Neuroscientists train a deep neural network to process sounds like humans do

19 April 2018



The Primary Auditory Cortex is highlighted in magenta, and has been known to interact with all areas highlighted on this neural map. Credit: Wikipedia.

Using a machine-learning system known as a deep neural network, MIT researchers have created the first model that can replicate human performance on auditory tasks such as identifying a musical genre.

This model, which consists of many layers of information-processing units that can be trained on huge volumes of data to perform specific tasks, was used by the researchers to shed light on how the [human brain](#) may be performing the same tasks.

"What these models give us, for the first time, is machine systems that can perform sensory tasks that matter to humans and that do so at human levels," says Josh McDermott, the Frederick A. and Carole J. Middleton Assistant Professor of Neuroscience in the Department of Brain and Cognitive Sciences at MIT and the senior author of the study. "Historically, this type of sensory processing has been difficult to understand, in part

because we haven't really had a very clear theoretical foundation and a good way to develop models of what might be going on."

The study, which appears in the April 19 issue of *Neuron*, also offers evidence that the human auditory cortex is arranged in a hierarchical organization, much like the visual cortex. In this type of arrangement, sensory information passes through successive stages of processing, with basic information processed earlier and more advanced features such as word meaning extracted in later stages.

MIT graduate student Alexander Kell and Stanford University Assistant Professor Daniel Yamins are the paper's lead authors. Other authors are former MIT visiting student Erica Shook and former MIT postdoc Sam Norman-Haignere.

Modeling the brain

When deep neural networks were first developed in the 1980s, neuroscientists hoped that such systems could be used to model the human [brain](#). However, computers from that era were not powerful enough to build models large enough to perform real-world tasks such as object recognition or speech recognition.

Over the past five years, advances in computing power and [neural network](#) technology have made it possible to use neural networks to perform difficult real-world tasks, and they have become the standard approach in many engineering applications. In parallel, some neuroscientists have revisited the possibility that these systems might be used to model the human brain.

"That's been an exciting opportunity for neuroscience, in that we can actually create systems that can do some of the things people can do, and we can then interrogate the models and compare them to the brain," Kell says.

The MIT researchers trained their neural network to perform two auditory tasks, one involving speech and the other involving music. For the speech task, the researchers gave the model thousands of two-second recordings of a person talking. The task was to identify the word in the middle of the clip. For the music task, the model was asked to identify the genre of a two-second clip of music. Each clip also included background noise to make the task more realistic (and more difficult).

After many thousands of examples, the model learned to perform the task just as accurately as a human listener.

"The idea is over time the model gets better and better at the task," Kell says. "The hope is that it's learning something general, so if you present a new sound that the model has never heard before, it will do well, and in practice that is often the case."

The model also tended to make mistakes on the same clips that humans made the most mistakes on.

The processing units that make up a neural network can be combined in a variety of ways, forming different architectures that affect the performance of the model.

The MIT team discovered that the best model for these two tasks was one that divided the processing into two sets of stages. The first set of stages was shared between tasks, but after that, it split into two branches for further analysis—one branch for the speech task, and one for the musical genre task.

Evidence for hierarchy

The researchers then used their model to explore a longstanding question about the structure of the auditory cortex: whether it is organized hierarchically.

In a hierarchical system, a series of brain regions performs different types of computation on sensory information as it flows through the system. It has been well documented that the visual cortex has this type of organization. Earlier regions, known as

the [primary visual cortex](#), respond to simple features such as color or orientation. Later stages enable more complex tasks such as object recognition.

However, it has been difficult to test whether this type of organization also exists in the auditory cortex, in part because there haven't been good models that can replicate human auditory behavior.

"We thought that if we could construct a model that could do some of the same things that people do, we might then be able to compare different stages of the model to different parts of the brain and get some evidence for whether those parts of the brain might be hierarchically organized," McDermott says.

The researchers found that in their model, basic features of sound such as frequency are easier to extract in the early stages. As information is processed and moves farther along the network, it becomes harder to extract frequency but easier to extract higher-level information such as words.

To see if the model stages might replicate how the human auditory cortex processes sound information, the researchers used functional magnetic resonance imaging (fMRI) to measure different regions of auditory cortex as the brain processes real-world sounds. They then compared the brain responses to the responses in the model when it processed the same sounds.

They found that the middle stages of the model corresponded best to activity in the primary auditory cortex, and later stages corresponded best to activity outside of the primary cortex. This provides evidence that the auditory [cortex](#) might be arranged in a hierarchical fashion, similar to the [visual cortex](#), the researchers say.

"What we see very clearly is a distinction between primary [auditory cortex](#) and everything else," McDermott says.

The authors now plan to develop models that can perform other types of auditory tasks, such as determining the location from which a particular sound came, to explore whether these tasks can be done by the pathways identified in this [model](#) or if

they require separate pathways, which could then be investigated in the brain.

Provided by Massachusetts Institute of Technology

APA citation: Neuroscientists train a deep neural network to process sounds like humans do (2018, April 19) retrieved 21 September 2020 from <https://techxplore.com/news/2018-04-neuroscientists-deep-neural-network-humans.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.