

NEWS

How artificial intelligence is shaking up animal behavior studies in autism

BY ALLA KATSNELSON

24 MAY 2021

Listen to this story:

<https://www.spectrumnews.org/wp-content/uploads/2021/05/audio-fb2154ea-bb22-4d3c-bfdf-fa9909591779-encodings.mp3>

Sam Golden never set out to develop machine-learning software that analyzes mouse behavior. But when he launched his lab in January 2019 to study the neurobiology of aggression, he and his tech-savvy postdoctoral fellow were quick to agree on what they didn't want to do.

"We just looked at each other and said, 'We don't want to score any more videos,'" says Golden, assistant professor of neuroscience at the University of Washington in Seattle.

Their invention — a package called **Simple Behavioral Analysis** (SimBA) — was born of a frustration familiar to many autism researchers. Studying behavior generally involves recording videos of rodents, flies, fish, birds, monkeys or other animals — and then 'scoring' them by watching hours of replays with a clicker in hand, counting and categorizing what the animals are doing: How many times did that mouse approach another? Or bury a marble in its bedding?

The past few years have seen a proliferation of open-source tools that help to automate that process. Many use machine learning to identify multiple points on an animal's body, such as the nose and key joints, and use those points to track the animal's poses. Researchers can statistically analyze these 'pose estimation' data to determine the animal's movements and autoscore its behavior.

Autoscore: Using new machine-learning tools, researchers can track how an animal moves and define the components movements of complex behaviors — say, rearing, grooming or fighting.

Photograph by M. Scott Brauer

But a new generation of tools — including SimBA; **Motion Sequencing**, or MoSeq, from **Sandeep Robert Datta**'s lab at Harvard University; and the **Mouse Action Recognition System** (MARS), from **David Anderson** and **Pietro Perona**'s labs at the California Institute of Technology in Pasadena — is taking the analysis a step further.

Using these newer algorithms, researchers can not only track how an animal moves, but define the component movements of more complex behaviors — say, rearing, grooming or fighting. And the programs can be trained to home in on a specific behavior, or they can be unleashed on video data with no prior training to pull out salient differences — in, for example, how male versus female mice fight or play.

For autism research, such software could be a major boon, scientists say. The algorithms could systematically compare social interactions among wildtype mice or rats with those of mice or rats missing different autism-related genes, Datta says. He is collaborating with autism researchers to conduct such studies. The tools could be used to zoom in on elements of social behavior shared across species. And they could also be used to rapidly **assess the behavioral effects** of drugs or other potential therapies for autism, as Datta's lab described last year.

“People are beginning to take movies of freely behaving animals and put them through deep networks to try to understand how animal behavior might be structured,” Datta says. “These technologies are really improving the ability of a wide range of scientists to ask all sorts of questions.”

Define, test, repeat: The new crop of automated tools might fix the lack of consistency between researchers in how they classify specific behaviors.

Photograph by M. Scott Brauer

Objective measures:

The new crop of automated tools might also solve a major problem that has plagued behavioral studies for decades: a lack of consistency in how researchers between labs, and even within a

single lab, classify specific behaviors.

For example, a social interaction between two mice might consist of several discrete behaviors, including posturing or approaching, sniffing or grooming the other animal. Two researchers studying the behavior might hold different assumptions about which of these components must be present to count the interaction during manual scoring. What's more, many behaviors exist on a gradient. What constitutes excessive grooming is a judgement call that two researchers might make differently.

"I think this will be their major contribution — transparency and reproducibility across labs," says **Cornelius Gross**, a neuroscientist and interim head of the European Molecular Biology Laboratory in Rome, Italy.

Because the algorithms need to quantitatively define the parameters of a particular behavior, their measures are objective. Researchers can compare the parameters they are using with those that their colleagues use to define the same behavior — or determine whether a difference in results stems from a behavioral difference or one in how the behavior is defined.

"It just becomes less of a subjective observation, and really something that is concrete and reproducible," Golden says.

The technology also makes it possible to correlate specific behaviors with brain activity, to a degree of precision that manual scoring doesn't allow, says **Ann Kennedy**, assistant professor of neuroscience at Northwestern University in Chicago. "We can see what the brain is doing and relate it to the animal's actions."

Researchers can record from the brain using multiple electrodes to capture the activity of multiple brain circuits, or they can use techniques such as optogenetics, turning specific circuits on and off, and then correlate the timing of the switch with second-to-second behavioral changes. In this way, researchers can pinpoint the brain activity that underlies a specific action — say, the initiation of a social interaction versus an aversion to one.

While Kennedy was a postdoctoral researcher in Anderson's lab at the California Institute of Technology in Pasadena, she and her colleagues used MARS to show that **brain activity patterns** in male mice differ depending on whether they are mounting females or other males.

"The two behaviors look indistinguishable," Kennedy says, "but the neural correlates of the behavior are different."

Spot check: SimBA includes "manual checkpoints" for researchers to review video clips of what the algorithm is predicting to be a particular behavior and assess its accuracy.

Photograph by Rick Dahms

Out of context:

Not everyone is eager to embrace the new tools, however.

Behavioral research in autism is especially fraught, says **Mu Yang**, assistant professor of neurobiology at Columbia University and director of the Mouse Neurobehavior Core, because most researchers have a poor handle on interpreting social behavior in different species. This makes it challenging to compare these behaviors with differences in social interaction among autistic people. Using automated tools to assess those behavioral changes might only further muddy scientists' interpretations.

"[The tools] give people the impression that we can now buy some plug-and-play stuff and, ta-da! Out will come autism-relevant phenotypes without anyone having to look at any behaviors," Yang says.

Other scientists trained in traditional ethology are quick to note that animal behaviors, including learning and communication, don't occur in a vacuum. Relying on machine learning to categorize why two mice interact in a particular way in a particular moment, for example, could, when taken out of context, lead researchers to misinterpret them, says **Jill Silverman**, associate professor of psychiatry and behavioral sciences at the University of California, Davis.

Automation can be useful for some simple behaviors, she says, but "you can't be a behaviorist and not watch behavior."

SimBA includes "manual checkpoints" for researchers to review video clips of what the algorithm is predicting to be a particular behavior and assess its accuracy, Golden says. "People are totally correct to be freaked out by [full automation]."

Yet for researchers who lack extensive experience parsing social behaviors — for example, interpreting whether one mouse is avoiding another because it's anxious or because it would simply rather sniff a piece of cardboard — such a check wouldn't necessarily help, Yang says. If someone inexperienced does the checking, she says, "It is the blind leading the blind."

For now, at least, manual scoring is likely to remain a mainstay. Behavioral analysis algorithms as a whole still have many technical limitations. Among them is the fact that researchers often have to retrain their algorithm whenever they change their experimental setup because strong visual differences make the elements the algorithm trained on less recognizable. And some algorithms do a poor job of tracking two similar-looking and closely interacting animals; they are prone to confuse

the points they are tracking, like two nuzzling noses.

Within a whisker: Some algorithms do a poor job of tracking two closely interacting animals and are prone to confuse the points they are tracking, like two nuzzling noses.

Photograph by M. Scott Brauer

Gross' lab uses SimBA and other AI tools, but he says he is not yet convinced that it's worth the investment to train students to use the programs. In his experience, the programs mix up two mice too frequently, especially under low-contrast lighting, he says. And picking through the videos to check for these sorts of errors is time consuming — almost as time consuming as scoring them by hand.

“A single error of swapping the identity of two animals during ‘crossover’ can ruin the [data for the] whole video,” he says.

As machine learning improves, those issues will get resolved, Golden says. In April, his lab released an updated version of SimBA that can classify behaviors not just from video but in real time, enabling researchers to precisely time when they switch on an optogenetically controlled circuit, for example. He is also working with a consortium of labs worldwide to standardize how the parameters for specific behaviors can be shared.

“The message that we are trying to send is, give it a chance,” he says. “It’s not as scary as it looks.”