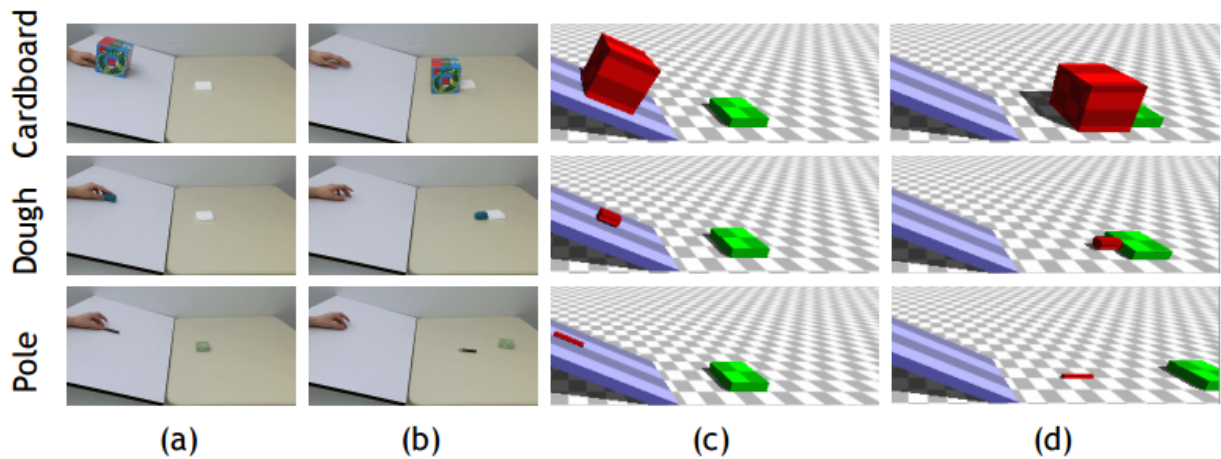# Computer model matches humans at predicting how objects move

January 5 2016, by Adam Conner-Simons

Researchers fed their physics prediction system videos of collisions, shown as screenshots in (a) and (b), which were then converted into simulations generated by the 3-D physics engine, shown in (c) and (d).

We humans take for granted our remarkable ability to predict things that happen around us. For example, consider Rube Goldberg machines: One of the reasons we enjoy them is because we can watch a chain-reaction of objects fall, roll, slide and collide, and anticipate what happens next.

But how do we do it? How do we effortlessly absorb enough information from the world to be able to react to our surroundings in real-time? And, as a computer scientist might then wonder, is this something that we can

teach machines?

That last question has recently been partially answered by researchers at MIT's Computer Science and Artificial Intelligence Lab (CSAIL), who have developed a computational model that is just as accurate as humans at predicting how objects move.

By training itself on real-world videos and using a "3-D physics engine" to simulate human intuition, the system—dubbed "Galileo"—can infer the [physical properties](#) of objects and predict the outcome of a variety of physical events.

While the researchers' paper focused on relatively simple experiments involving ramps and collisions, they say that the model's ability to generalize its findings and continuously improve itself means that it could readily predict a range of actions.

"From a ramp scenario, for example, Galileo can infer the density of an [object](#) and then predict if it can float," says postdoc Ilker Yildirim, who was lead author alongside CSAIL PhD student Jiajun Wu. "This is just the first step in imbuing computers with a deeper understanding of dynamic scenes as they unfold."

The paper, which was presented this past month at the Conference on Neural Information Processing Systems (NIPS) in Montreal, was co-authored by postdoc Joseph Lim and professor William Freeman, as well as professor Joshua Tenenbaum from the Department of Brain and Cognitive Sciences.

How they did it

Recent research in neuroscience suggests that in order for humans to understand a scene and predict events within it, our brains rely on a

mental "physics engine" consisting of detailed but noisy knowledge about the physical laws that govern objects and the larger world.

Using the human framework to develop their model, the researchers first trained Galileo on a set of 150 videos that depict physical events involving objects of 15 different materials, from cardboard and foam to metal and rubber. This training allowed the model to generate a dataset of objects and their various physical properties, including shape, volume, mass, friction, and position in space.

From there, the team fed the model information from Bullet, a 3-D physics engine often used to create special effects for movies and video games. By inputting the setup of a given scene and then physically simulating it forward in time, Bullet serves as a reality check against Galileo's hypotheses.

Finally, the team developed deep-learning algorithms that allow the model to teach itself to further improve its predictions to the point that, by the very first frame of a video, Galileo can recognize the objects in the scene, infer the properties of the objects, and determine how these objects will interact with one another.

"Humans learn physical properties by actively interacting with the world, but for our computers this is tricky because there is no training data," says Abhinav Gupta, an assistant professor of computer science at Carnegie Mellon University. "This paper solves this problem in a beautiful manner, by combining deep-learning convolutional networks with classical AI ideas like simulation engines."

Human vs. machine

To assess Galileo's predictive powers, the team pitted it against human subjects to predict a series of simulations (including one that has an

online demo).

In one, users see a series of object collisions, and then another video that stops at the moment of collision. The users are then asked to label how far they think the object will move.

"The scenario seems simple, but there are many different physical forces that make it difficult for a computer model to predict, from the objects' relative mass and elasticity to gravity and the friction between surface and object," Yildirim says. "Where humans learn to make such judgments intuitively, we essentially had to teach the system each of these properties and how they impact each other collectively."

In another simulation, users first see a collision involving a 20-degree inclined ramp, and then are shown the first frame of a video with a 10-degree ramp and asked to predict whether the object will slide down the surface.

"Interestingly, both the computer model and human subjects perform this task at chance and have a bias at saying that the object will move," Yildirim says. "This suggests not only that humans and computers make similar errors, but provides further evidence that human scene understanding can be best described as probabilistic simulation."

What's next

The team members say that they plan to extend the research to more complex scenarios involving fluids, springs, and other materials. Continued progress in this line of work, they say, could lead to direct applications in robotics and [artificial intelligence](link).

"Imagine a robot that can readily adapt to an extreme physical event like a tornado or an earthquake," Lim says. "Ultimately, our goal is to create

flexible models that can assist humans in settings like that, where there is significant uncertainty."

   More information: Galileo: Perceiving Physical Object Properties by Integrating a Physics Engine with Deep Learning. www.mit.edu/~ilkery/papers/phys_nips.pdf

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: Computer model matches humans at predicting how objects move (2016, January 5) retrieved 9 April 2024 from https://phys.org/news/2016-01-humans.html