

## Technology to make self-driving cars, robotics, and other applications understand the 3-D world

October 22 2019, by Adam Conner-Simons



At left, EdgeConv, a method developed at MIT, successfully finds meaningful parts of 3D shapes, like the surface of a table, wings of an airplane, and wheels of a skateboard. At right is the ground-truth comparison. Credit: Massachusetts Institute of Technology



If you've ever seen a self-driving car in the wild, you might wonder about that spinning cylinder on top of it.

It's a "lidar sensor," and it's what allows the car to navigate the world. By sending out pulses of infrared light and measuring the time it takes for them to bounce off objects, the sensor creates a "point cloud" that builds a 3-D snapshot of the car's surroundings.

Making sense of raw point-cloud data is difficult, and before the age of machine learning it traditionally required highly trained engineers to tediously specify which qualities they wanted to capture by hand. But in a new series of papers out of MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL), researchers show that they can use deep learning to automatically process point clouds for a wide range of 3-D-imaging applications.

"In computer vision and machine learning today, 90 percent of the advances deal only with <u>two-dimensional images</u>," says MIT Professor Justin Solomon, who was senior author of the new series of papers spearheaded by Ph.D. student Yue Wang. "Our work aims to address a fundamental need to better represent the 3-D world, with application not just in autonomous driving, but any field that requires understanding 3-D shapes."

Most previous approaches haven't been especially successful at capturing the patterns from data that are needed to get <u>meaningful information</u> out of a bunch of 3-D points in space. But in one of the team's papers, they showed that their "EdgeConv" method of analyzing point clouds using a type of neural network called a dynamic graph convolutional neural network allowed them to classify and segment individual objects.

"By building 'graphs' of neighboring points, the algorithm can capture hierarchical patterns and therefore infer multiple types of generic



information that can be used by a myriad of downstream tasks," says Wadim Kehl, a machine learning scientist at Toyota Research Institute who was not involved in the work.

In addition to developing EdgeConv, the team also explored other specific aspects of point-cloud processing. For example, one challenge is that most sensors change perspectives as they move around the 3-D world; every time we take a new scan of the same object, its position may be different than the last time we saw it. To merge multiple point clouds together into a single detailed view of the world, you need to align multiple 3-D points in a process called "registration."

Registration is vital for many forms of imaging, from satellite data to medical procedures. For example, when a doctor has to take multiple magnetic resonance imaging scans of a patient over time, registration is what makes it possible to align the scans to see what's changed.

"Registration is what allows us to integrate 3-D data from different sources into a common coordinate system," says Wang. "Without it, we wouldn't actually be able to get as meaningful information from all these methods that have been developed."

Solomon and Wang's second paper demonstrates a new registration algorithm called "Deep Closest Point" (DCP) that was shown to better find a point cloud's distinguishing patterns, points, and edges (known as "local features") in order to align it with other point <u>clouds</u>. This is especially important for such tasks as enabling self-driving cars to situate themselves in a scene ("localization"), as well as for robotic hands to locate and grasp individual objects.

One limitation of DCP is that it assumes we can see an entire shape instead of just one side. This means it can't handle the more difficult task of aligning partial views of shapes (known as "partial-to-partial



registration"). As a result, in a third paper the researchers presented an improved algorithm for this task that they call the Partial Registration Network (PRNet).

Solomon says that existing 3-D data tends to be "quite messy and unstructured compared to 2-D images and photographs." His team sought to figure out how to get meaningful information out of all that disorganized 3-D data without the controlled environment that a lot of machine learning technologies now require.

A key observation behind the success of DCP and PRNet is the idea that a critical aspect of point-cloud processing is context. The geometric features on point cloud A that suggest the best ways to align it to point cloud B may be different from the features needed to align it to point cloud C. For example, in partial registration, an interesting part of a shape in one point cloud may not be visible in the other—making it useless for registration.

Wang says that the team's tools have already been deployed by many researchers in the computer vision community and beyond. Even physicists are using them for an application the CSAIL team had never considered: particle physics.

Moving forward, the researchers hope to use the algorithms on realworld data, including data gathered from self-driving cars. Wang says they also plan to explore the potential of training their systems using selfsupervised learning, to minimize the amount of human annotation needed.

**More information:** Dynamic Graph CNN for Learning on Point Clouds: Dynamic Graph CNN for Learning on Point Clouds

Deep Closest Point: Learning Representations for Point Cloud



Registration: arxiv.org/abs/1905.03304

PRNet: Self-Supervised Learning for Partial-to-Partial Registration: nips.cc/Conferences/2019/Schedule?showEvent=13934

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: Technology to make self-driving cars, robotics, and other applications understand the 3-D world (2019, October 22) retrieved 28 April 2024 from <u>https://techxplore.com/news/2019-10-technology-self-driving-cars-robotics-applications.html</u>

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.