

Waymo's AI Content Search tool lets engineers quickly find objects in driving records

Kyle Wiggers

@Kyle_L_Wiggers

February 6, 2020 10:30 AM



Waymo's fully self-driving Jaguar I-PACE electric SUV
Image Credit: Waymo

Were you unable to attend Transform 2022? Check out all of the summit sessions in our on-demand library now! [Watch here](#).

AI is the method by which self-driving cars perceive joggers, cyclists, traffic lights, road signs, trees, shrubs, and more, and it informs the way in which they choose to behave when encountered with those signals. The vehicles in [Waymo's](#) fleet aren't an exception to the rule — they tap AI to make real-time driving decisions, in part by matching obstacles spotted by their onboard sensors to the billions of objects in the Alphabet company's database.

Large data sets are invaluable in the autonomous driving domain because they enable the underpinning AI to self-improve. But it's been historically tough for engineers to surface samples within those sets without investing time and manual effort. That's why Waymo developed what it

calls Content Search, which draws on tech similar to that which powers Google Photos and Google Image Search to let data scientists quickly locate almost any object in Waymo's driving history and logs.

Waymo previously collaborated with Alphabet's DeepMind on AI techniques inspired by evolutionary biology. The latter's PBT (Population Based Training), which starts with multiple machine learning models and replaces underperforming members with "offspring," managed to reduce false positives by 24% in pedestrian, bicyclist, and motorcyclist recognition tasks while cutting training time and computational resources in half. In point of fact, following a pilot study, PBT was integrated directly with Waymo's technical infrastructure, enabling researchers from across the company to apply it with a button click.

Mining for data

Waymo notes that the Waymo Driver — its stack of driverless vehicle perception and decision-making technologies — is fleet-aware, so that everything one car learns can be shared among the rest. (The Chrysler Pacifica minivans like those deployed as part of Waymo One, Waymo's commercial robo-taxi service, exchange information about hazards and route changes via a set of dual wireless modems.) That's especially useful in light of emerging trends in transportation, such as the recent rise in the popularity of electric scooters and bicycles.



"Waymo Driver regularly encounters new forms of transportation," explained Waymo in a blog post. "So we want to continually train our system to ensure that we can not only distinguish between a vehicle and a cyclist, but also between a pedestrian and a person on a scooter."

In the past, Waymo researchers had to rely on heuristic methods to find distinct samples in Waymo's logs, many of which simply parsed data based on various features (i.e., an object's estimated speed and height). For instance, to locate examples of people riding on scooters, a Waymo engineer might have specified that they wished to see objects of a certain height traveling between 0 and 20 miles per hour. But the results were often too broad.

“Oversize”



ADVERTISEMENT

By contrast, Content Search approaches this sort of data mining task as a search problem. It indexes catalogs of data, conducting similarity searches to pinpoint objects in “ultra-fine-grained categories” and identify text, which allows it to find items similar to logged objects by running image comparison queries. Given an image of a cactus, say — whether a cactus from driving history or even a drawing of a cactus — Content Search returns instances where Waymo’s self-driving vehicles observed similar-looking objects.

Machine learning inside

Content Search works its magic by converting every object in Waymo’s driving logs — whether a park bench, a trash can near the side of the road, or a moving object — into embeddings, or mathematical representations derived from attributes. This lets the tool rank objects by how similar they are to each other, akin to the process employed by Google’s [real-time embedding similarities matching service](#). Better still, it enables Content Search to compare queries against images in the data logs and locate similar objects in a matter of seconds.

Debris



Generalizing and detecting even a single class of object is no walk in the park, Waymo points out — objects vary in shape, form, and type and range from plastic bags to tire scraps to cardboard boxes and lost pairs of pants. The company's engineers train a range of AI with diverse examples to help classify items familiar and unfamiliar, and they use a backend for Content Search that employs a categorical AI model to understand whether an object category is present in an image. Plus, they maintain a state-of-the-art optical character recognition model to annotate driving logs based on text and words found in a scene, enabling Content Search to comprehend things like road signs, emergency vehicles, and other cars and trucks with signage (like the "oversized" notice on a large truck).

This allows Waymo's engineers to perform niche searches on objects that share a particular trait, such as the make and model of a car or even specific breeds of dogs.

"With Content Search, we're able to automatically annotate ... objects in our driving history which in turn has exponentially increased the speed and quality of data we send for labeling," wrote Waymo. "The ability to accelerate labeling has contributed to many improvements across our system, from detecting school buses with children about to step onto the sidewalk or people riding electric scooters to a cat or a dog crossing a street. As Waymo expands to more cities, we'll continue to encounter new objects and scenarios."