Amazon’s Larry Davis on the past and future of computer vision research.

By [Larry Hardesty](https://www.amazon.science/author/larry-hardesty)

June 11, 2020

The IEEE Conference on Computer Vision and Pattern Recognition ([CVPR](https://www.amazon.science/conferences-and-events/cvpr-2020)), which will be held virtually this year and starts next week, is the major conference in the field of computer vision. But when Larry Davis, senior principal scientist for Amazon Fashion, started attending it, “computer vision” wasn’t even part of its title.

“I think 1981 was my first time having a formal role at the conference,” Davis says. “Back then the meeting was called Pattern Recognition and Image Processing. My advisor [Azriel Rosenfeld] was pretty much regarded as the founder of the field of computer vision, but when he founded it, he called it ‘picture processing’. Then he and a couple other senior people decided to rename it computer vision, and they changed the name of the conference to CVPR.”

In his four decades attending CVPR, Davis has witnessed several waves of change sweep through the computer vision field.

Larry Davis, senior principal scientist for Amazon Fashion.

“The 1980s was the decade of what we would call segmentation,” Davis says. “How do you take a picture and break it into parts? One of the motivations was computational — trying to reduce the combinatorics of finding objects in pictures. What’s an object? You might say it’s some connected set of pixels. There are a lot of connected subsets of pixels in images. You can’t begin to look for objects by enumerating the connected subsets. But suppose you could break the image up into 100 pieces via color and texture segmentation. Now you can begin to think of a more-or-less brute-force algorithm that would look at combinations of pieces and ask, ‘Is this the right shape for a dog?’

“Alternatively, you could try to outline objects. Perceptual psychologists will tell you that most of the information that people perceive from images is from the boundaries of objects or the boundaries between the objects and background. So people tried for decades to build edge detectors that would enclose objects. That kind of died out. People weren’t making much progress.

“In the ’90s there was an interest in geometry — multiview geometry, how do you build 3-D models from multiple images from a camera moving through the environment. There was a ton of great work on that.

“In the 2000s, computers became fast enough, mass storage became cheap enough, that people were doing video, so there was a lot of interest in video surveillance. Once the social-media companies took hold, then the emphasis in video was more on consumer videos, being able to caption them, summarize them visually, index into them, all kinds of things.

“And then, the last seven or eight years it’s been all deep learning for all these problems. And that’s when the field took off in terms of size and sophistication.”

**Paradigm shift**

As it has in other areas of AI, deep learning has revolutionized computer vision.

“Instead of sitting around trying to design new feature detectors, you design new architectures that learn the features,” Davis says.

For researchers of Davis’s generation, that sometimes meant the abandonment of long-standing research programs.

“I saw in a number of my colleagues a resistance to the deep-learning wave,” Davis says. “I met [deep-learning pioneer] [Geoff Hinton](https://amturing.acm.org/award_winners/hinton_4791679.cfm%22%20%5Ct%20%22_blank) and a researcher then at Hopkins named [Terry Sejnowski](https://www.salk.edu/scientist/terrence-sejnowski/%22%20%5Ct%20%22_blank) — he was more a biomedical type who was interested in neural computing — when they were young. I always thought they were brilliant. It wasn’t what I did, but I was willing to believe them. So it never really bothered me to vacate classical computer vision and just parachute into deep-learning territory. And in any event, any faculty members who didn’t do it were dragged in by their students, because they certainly knew what kind of skills they needed to develop in order to be successful.”

For all of deep learning’s successes, however, some researchers have begun to question whether we can expect to get much more mileage out of it.

“I think we might have hit a wall on some problems — for example, there have been no dramatic improvements recently in the ability to detect objects in pictures,” Davis says. “If you want to build the world’s ultimate flower detector, that never makes a mistake, that sees every flower from any angle, even when only a very small part of the flower is visible, or its image is very small, you realize that might never happen. Even if the problem is just the unavailability of training data to cover all edge cases — if they could ever be known — the cost of doing this for all object types, maybe even just for flowers, would be prohibitive today.”

It never really bothered me to vacate classical computer vision and just parachute into deep-learning territory. ... I think the field attracts people who are not challenged by change.

Larry Davis

Davis, however, remains optimistic about the continued applicability of deep-learning techniques. While object recognition may be running up against its limits, he says, “the interesting thing is that there are still so many more problems that you can successfully approach. They’re just different problems, and even the current object detectors are good enough to solve lots of important problems.”

Within Amazon Fashion, for instance, “there are several different things going on that involve different types of deep-learning architecture, and they all have some scientific creativity, something new that somebody’s developed,” Davis says.

Indeed, three of the ten [Amazon papers](https://www.amazon.science/conferences-and-events/cvpr-2020%22%20%5Ct%20%22_blank) at CVPR this year concern ways in which computer vision can help improve customers’ experience while [shopping for clothing online](https://www.amazon.science/blog/how-computer-vision-will-help-amazon-customers-shop-online%22%20%5Ct%20%22_blank).

“The number of deep-learning problems whose solution would improve our customers’ shopping experience is enormous, and we’ve expanded our science team to tackle more and more of them,” Davis says. “We’re always looking for more ML scientists excited about fashion to join us.”

Still, if some new, more powerful artificial-intelligence paradigm arises, Davis will remain no more attached to deep learning than he was to pixel clustering and boundary detection.

“In computer science generally, things change very rapidly,” he says. “You can’t hold on to old stuff. I think the field attracts people who are not challenged by change.”