UC SANTA BARBARA

THE Current

June 18, 2019
Sonia Fernandez

The Intersection of Vision and Language

Nine thousand two hundred artificial intelligence researchers. Five thousand one hundred sixty-five research papers submitted, of which only 1,300 were accepted. One Best Student Paper.

And the award went to: UC Santa Barbara computer science doctoral student Xin Wang. His student paper, "Reinforced Cross-Modal Matching and Self-Supervised Imitation Learning for Vision-Language Navigation," emerged No. 1 in his category at the 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), one of the most important conferences in the field of computer science today. The awards event took place in Long Beach, on Tuesday, June 18.

"Xin started working with me in 2017 as a Ph.D. student on topics related to language and vision research," said William Wang, assistant professor in the Department of Computer Science at UCSB's College of Engineering. "Since March 2018, we have studied the vision-language navigation problem: How do you use verbal instructions to teach robots to target destinations, without access to a map?"

A robot that can navigate based on vision and perform tasks under instructions in ordinary language sounds like science fiction, but it's closer to reality than you might think.

“I have been working on the intersection of vision and language for a while,” Xin Wang said, “and I believe one of the big moves of AI is to have robots interact with the visual and physical world, especially via natural language.

“Vision-language navigation can enable many practical applications, for example, in-home robots,” he continued. “Everyone, especially the disabled, can ask the robot to perform some certain tasks, for example ‘Go to the kitchen and bring a cup of water to me.’ Vision-language navigation serves as a basic task to understand both visual scenes and natural language instructions, as well as perform physical actions to fulfill high-level jobs.”

But what comes so naturally to people — navigating through a scene and performing tasks using real-time environmental cues and human concepts — requires a flexible, yet elegant framework in which robots can learn to connect the data they pick up to the meaning of the instructions they receive. It’s an iterative process, requiring reinforcement, feedback and adaptation. Xin Wang’s work, according to his advisor, “introduces several innovative ideas to improve generalization of the vision-language navigation algorithm.”

“This breakthrough in language and vision research will enable robots to better assist humans in many daily and special routines, including home cleaning and maintenance, item finding and retrieving, remote control, assisting blind people, disaster relief, et cetera,” William Wang said. “It has the potential to influence millions of people and improve the life quality of humans, including freeing humans from tedious household tasks, so that we can make time for creative activities.”

The research for the 2019 CVPR Best Student Paper, which builds upon previous collaborations at UCSB on model-based and model-free reinforcement learning with Ph.D. student Wenhan Xiong, was conducted in the summer of 2018 during an internship at Microsoft Research (MSR), under mentors Qiuyuan Huang, Asli Celikyilmaz, Jianfeng Gao and Lei Zhang.

According to the CVPR Best Paper committee, “Visual navigation is an important area of computer vision — this paper makes advances in vision-language navigation. Building on previous work in this area, this paper demonstrates exciting results based on self-imitation learning within a cross-modal setting.”

“We would like to congratulate Xin Wang and William Wang on this extremely prestigious recognition for their work in an area that is at the very leading edge of

computer science," said College of Engineering Dean Rod Alferness. "We are delighted that they are part of the UCSB community."

"I am really honored to receive this award," Xin Wang said. "I would like to sincerely thank my advisors William and Yuan-Fang, and the MSR collaborators for their strong support and valuable guidance. As for the future, I hope that more and more researchers can work on this exciting and necessary research direction, toward more practical and interactive robots that bridge vision and language for humans. I will certainly devote myself to making it happen."

Tags
Artificial Intelligence

About UC Santa Barbara

The University of California, Santa Barbara is a leading research institution that also provides a comprehensive liberal arts learning experience. Our academic community of faculty, students, and staff is characterized by a culture of interdisciplinary collaboration that is responsive to the needs of our multicultural and global society. All of this takes place within a living and learning environment like no other, as we draw inspiration from the beauty and resources of our extraordinary location at the edge of the Pacific Ocean.