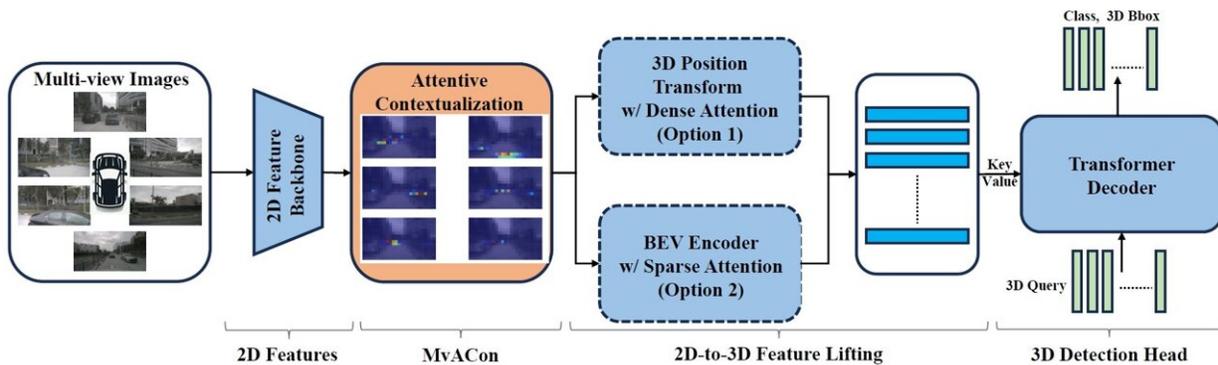


New technique improves AI ability to map 3D space with 2D cameras

June 13 2024, by Matt Shipman



MvACon Credit: Liu et al.

Researchers have developed a technique that allows artificial intelligence (AI) programs to better map three-dimensional spaces using two-dimensional images captured by multiple cameras. Because the technique works effectively with limited computational resources, it holds promise for improving the navigation of autonomous vehicles.

"Most autonomous vehicles use powerful AI programs called vision transformers to take 2D images from multiple cameras and create a representation of the 3D space around the vehicle," says Tianfu Wu, corresponding author of a paper on the work and an associate professor of electrical and computer engineering at North Carolina State University. "However, while each of these AI programs takes a different

approach, there is still substantial room for improvement.

"Our technique, called Multi-View Attentive Contextualization (MvACon), is a plug-and-play supplement that can be used in conjunction with these existing vision transformer AIs to improve their ability to map 3D spaces," Wu says. "The vision transformers aren't getting any additional data from their cameras, they're just able to make better use of the data."

MvACon effectively works by modifying an approach called Patch-to-Cluster attention (PaCa), which Wu and his collaborators [released last year](#). PaCa allows transformer AIs to more efficiently and effectively identify objects in an image.

"The key advance here is applying what we demonstrated with PaCa to the challenge of mapping 3D space using multiple cameras," Wu says.

To test the performance of MvACon, the researchers used it in conjunction with three leading vision transformers—BEVFormer, the BEVFormer DFA3D variant, and PETR. In each case, the vision transformers were collecting 2D images from six different cameras. In all three instances, MvACon significantly improved the performance of each vision transformer.

"Performance was particularly improved when it came to locating objects, as well as the speed and orientation of those objects," says Wu. "And the increase in computational demand of adding MvACon to the vision transformers was almost negligible.

"Our next steps include testing MvACon against additional benchmark datasets, as well as testing it against actual video input from autonomous vehicles. If MvACon continues to outperform the existing [vision](#) transformers, we're optimistic that it will be adopted for widespread

use."

The paper, "[Multi-View Attentive Contextualization for Multi-View 3D Object Detection](#)," will be presented June 20 at the IEEE/CVF Conference on Computer Vision and Pattern Recognition, being held in Seattle, Wash.

The first author of the paper is Xianpeng Liu, a recent Ph.D. graduate of NC State. The paper was co-authored by Ce Zheng and Chen Chen of the University of Central Florida; Ming Qian and Nan Xue of the Ant Group; and Zhebin Zhang and Chen Li of the OPPO U.S. Research Center.

More information: Xianpeng Liu et al, [Multi-View Attentive Contextualization for Multi-View 3D Object Detection](#). (2024).

Provided by North Carolina State University

Citation: New technique improves AI ability to map 3D space with 2D cameras (2024, June 13) retrieved 29 June 2024 from <https://techxplore.com/news/2024-06-technique-ai-ability-3d-space.html>

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.