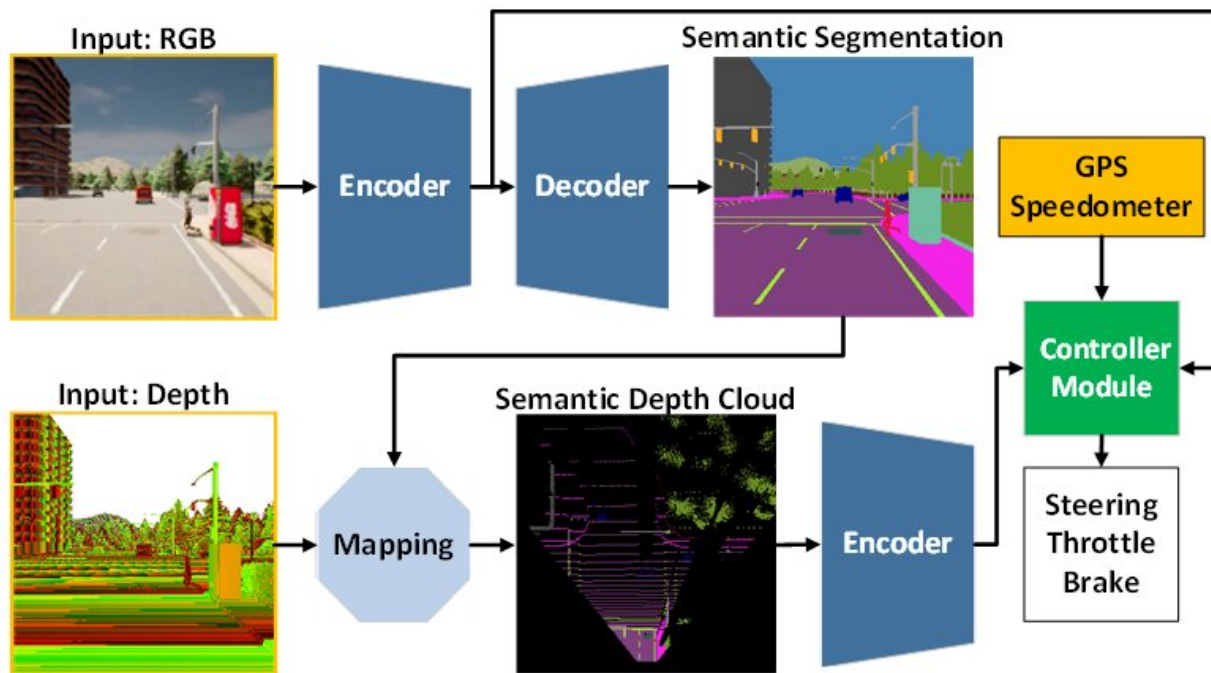# Researchers develop an AI model for autonomous driving

September 12 2022

The AI model architecture is composed of the perception module (blue) and the controller module (green). The perception module is responsible for perceiving the environment based on the observation data provided by an RGBD camera. Meanwhile, the controller module is responsible for decoding the extracted information to estimate the degree of steering, throttle, and braking. Credit: Toyohashi University of Technology

A research team consisting of Oskar Natan, a Ph.D. student, and his

supervisor, Professor Jun Miura, who are affiliated with the Active Intelligent System Laboratory (AISL), Department of Computer Science Engineering, Toyohashi University of Technology, has developed an AI model that can handle perception and control simultaneously for an autonomous driving vehicle.

The AI model perceives the environment by completing several vision tasks while driving the vehicle following a sequence of route points. Moreover, the AI model can drive the vehicle safely in diverse environmental conditions under various scenarios. Evaluated under point-to-point navigation tasks, the AI model achieves the best drivability of certain recent models in a standard simulation environment.

Autonomous driving is a complex system consisting of several subsystems that handle multiple perception and control tasks. However, deploying multiple task-specific modules is costly and inefficient, as numerous configurations are still needed to form an integrated modular system.

Furthermore, the integration process can lead to information loss as many parameters are adjusted manually. With rapid deep learning research, this issue can be tackled by training a single AI model with end-to-end and multi-task manners. Thus, the model can provide navigational controls solely based on the observations provided by a set of sensors. As manual configuration is no longer needed, the model can manage the information all by itself.

The challenge that remains for an end-to-end model is how to extract useful information so that the controller can estimate the navigational controls properly. This can be solved by providing a lot of data to the perception module to better perceive the surrounding environment. In addition, a sensor fusion technique can be used to enhance performance as it fuses different sensors to capture various data aspects.

However, a huge computation load is inevitable as a bigger model is needed to process more data. Moreover, a data preprocessing technique is necessary as varying sensors often come with different data modalities. Furthermore, imbalance learning during the training process could be another issue since the model performs both perception and control tasks simultaneously.
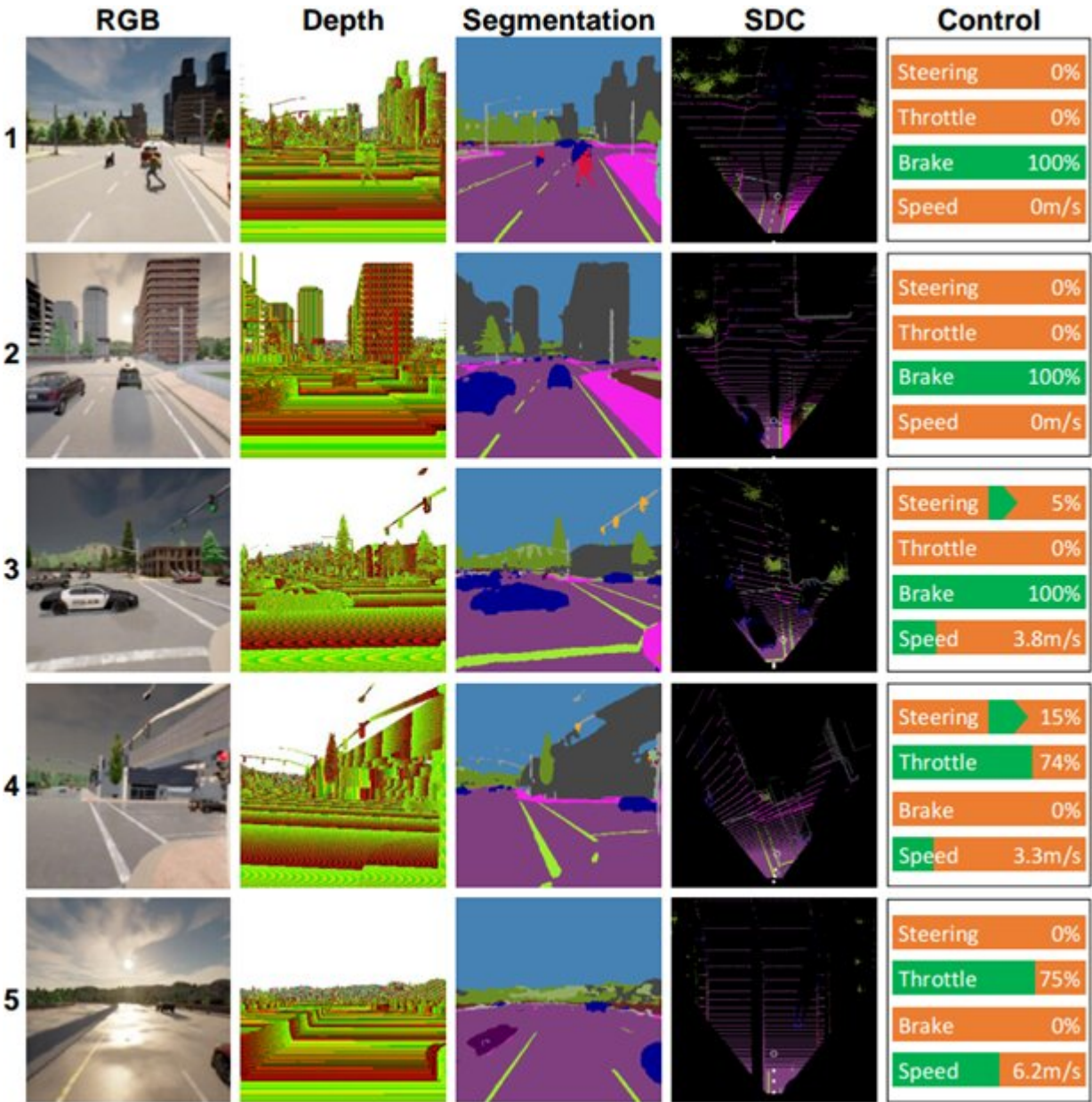
Some driving records made by the AI model. Columns (from left to right): color image, depth image, semantic segmentation result, birds-eye-view (BEV) map, control command. The weather and time for each scene are as follows: (1) clear noon, (2) cloudy sunset, (3) mid rainy noon, (4) hard rain sunset, (5) wet sunset. Credit: Toyohashi University of Technology

In order to answer those challenges, the team propose an AI model trained with end-to-end and multi-task manners. The model is made of two main modules, namely perception and controller modules. The perception phase begins by processing RGB images and depth maps provided by a single RGBD camera.

Then, the information extracted from the perception module along with vehicle speed measurement and route point coordinates are decoded by the controller module to estimate the navigational controls. So as to ensure that all tasks can be performed equally, the team employs an algorithm called modified gradient normalization (MGN) to balance the learning signal during the training process.

The team considers imitation learning as it allows the model to learn from a large-scale dataset to match a near-human standard. Furthermore, the team designed the model to use a smaller number of parameters than others to reduce the computational load and accelerate the inference on a device with limited resources.

Based on the experimental result in a standard autonomous driving simulator, CARLA, it is revealed that fusing RGB images and depth maps to form a birds-eye-view (BEV) semantic map can boost the overall performance. As the perception module has better overall understanding of the scene, the controller module can leverage useful information to estimate the navigational controls properly. Furthermore,

the team states that the proposed model is preferable for deployment as it achieves better drivability with fewer parameters than other models.

The research was published in *IEEE Transactions on Intelligent Vehicles*, and the team is currently working on modifications and improvements to the model so as to tackle several issues when driving in poor illumination conditions, such as at night, in heavy rain, etc. As a hypothesis, the team believes that adding a sensor that is unaffected by changes in brightness or illumination, such as LiDAR, will improve the model's scene understanding capabilities and result in better drivability. Another future task is to apply the proposed model to autonomous driving in the real world.

  More information: Oskar Natan et al, End-to-end Autonomous Driving with Semantic Depth Cloud Mapping and Multi-agent, *IEEE Transactions on Intelligent Vehicles* (2022). DOI: 10.1109/TIV.2022.3185303

Provided by Toyohashi University of Technology

Citation: Researchers develop an AI model for autonomous driving (2022, September 12) retrieved 16 August 2024 from https://techxplore.com/news/2022-09-ai-autonomous.html