Training and Development of Perception and Language Models for Autonomous Driving

About DEVCOM ARL:

DEVCOM ARL, as an integral part of the Army Futures Command, is the Army's foundational research laboratory focused on operationalizing science to ensure overmatch in any future conflict. DEVCOM ARL shapes future concepts with scientific research and knowledge and delivers technology for modernization solutions to win in the future operating environment.

RESEARCH LOCATION: Aberdeen Proving Ground, MD

PROJECT DESCRIPTION:

This research will focus on training and developing perception and language models for autonomous driving. It may include training perception models for single modality including LiDAR point clouds and images, and then apply data fusion technology to combine the information from multi-modality to provide better situation awareness of battlefield for unmanned Ground Vehicle to achieve autonomous driving. It may also include training the language pipeline to allow human control the UGV by using voice commands. The computationally expensive Deep Learning perception and language models will be trained in HPC using multiple GPUs. This research will directly support the ARL Artificial Intelligence for Maneuver and Mobility (AIMM) Essential Research Project.

The summer intern will be an integral part of the Resource Constraint Adaptive Computing (RCAC) team. The summer effort will begin by learning about the Deep Learning (DL) framework in HPC and understand DL perception and language models used for autonomous driving.

The summer intern will learn how to train DL models with single training node or with multiple training nodes across HPC. This will involve modifying the source code to support local training and distributed training in HPC, and then tuning training parameters for local and distributed training in order to achieve comparable performance. Multiple DL models need to be trained for different dataset including LiDAR point clouds, RGB images, and languages.

A state-of-the-art fusion model will be applied to combine various perception algorithms and provide better information for decision making. The source code needs to be rewritten to include the fusion model, and the combined model will be trained in HPC. Multiple language models will be trained and compared. The trained model can be evaluated in UGV husky or at least evaluated with collected rosbags.

The summer project will provide the intern the opportunity to learn the state-of-the-art Deep Learning (DL) perception and language models trained in HPC with multiple GPUs and use and enhance the skills learned through their academics. In addition to the HIP mini symposium, the intern will have the opportunity to present the summer research findings to ARL Summer Intern Symposium.

Summer interns at the ARL research in a collaborative workspace. This provides the students opportunities to interact, collaborate with each other and share their experiences. Mentors work with students regularly in this workspace, and hence students learn from other mentors as well as their own. Most importantly, this will provide an opportunity for the student as a potential DoD employee to experience DoD related research and development.

Intern's Project Schedule follows:

Week 1-2: Get up to speed on what the team has already done and get HPC account, HPCMP token, and ARL badge.

Weeks 3-4: Perform background research by reading papers to understand the perception and language models, and setup Conda environment in HPC including all necessary packages and datasets. Begin to train perception and language models in single training node in HPC.

Weeks 5-7: Expand the model training from single node to multiple nodes in HPC by modifying the source codes. Fine tune the training parameters of distributed training to achieve comparable accuracy with single node training while significantly reduce the training time.

Weeks 8-9: Train a data fusion model to combine multiple perception models together to achieve better situation awareness of the environment. Integrate and evaluate the perception and language models on UGVs if possible.

Week 10: Finalize briefing report, poster, and presentation.

ANTICIPATED START DATE:

June 2025 – Exact start dates will be determined at the time of selection and in coordination with the selected candidate.

QUALIFICATIONS:

Prospective candidates will major in computer science, computer engineer, electrical engineer, or other related engineering. The candidate has experience in administering Linux system and using Python, C++, and software repositories. Knowledge and experience of Deep Learning technology is a plus. Willing to learn the state-of-the-art DL technology is a big plus.

ACADEMIC LEVEL:

Degree received within the last 60 months or currently pursuing:

- Bachelor's
- Master's
- Doctoral

DISCIPLINE NEEDED:

- Computer, Information, and Data Science
- Engineering
- Science & Engineering-related