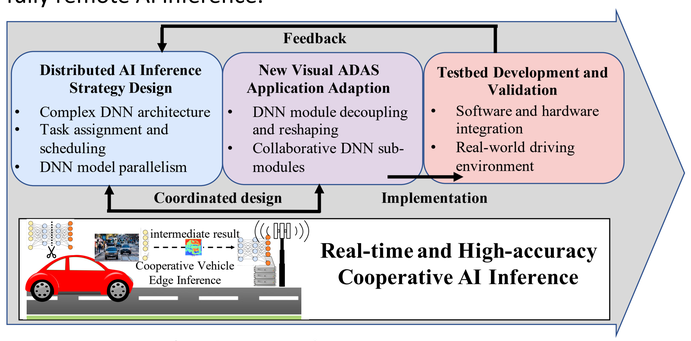
Cooperative AI Inference in Vehicular Edge Networks for Advanced Driver-Assistance Systems

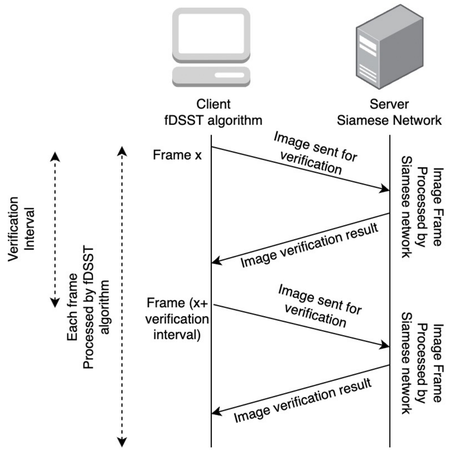
Shen-Shyang Ho (Rowan University), Haibin Ling (Stony Brook University), Jungme Park (Kettering University), Yu Wang (Temple University)

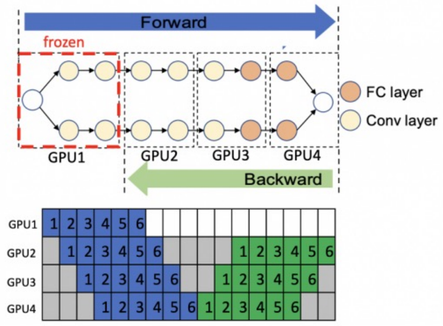
Project Overview:

Increasing driving safety for a wide range of commercial vehicles is one of the fundamental challenges in intelligent transportation systems. Recent Artificial Intelligence (AI) methodologies, especially Deep Neural Networks (DNNs), have led to superior performance in driving-assistant applications. However, the **high computation expense of DNN-based solutions often limits their deployment in the Advanced Driver-Assistance Systems (ADAS) of the majority of vehicles due to insufficient onboard processing capability.** To address this challenge, the overall objective of this project is to explore cooperative AI inference on ADAS with the emerging vehicular edge computing paradigm. Cooperative AI inference decomposes a DNN-based solution into a set of layers and distributes layers between the vehicle and server. This way it brings significant advantages over naive edge computation offloading that offloads raw sensory data to the edge server for fully remote AI inference.



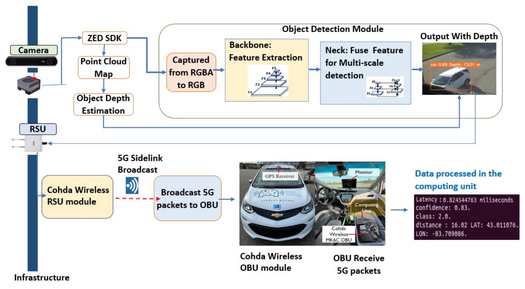
Research Highlights (Technical Approach, Key Innovations, New (23-24) Contributions)



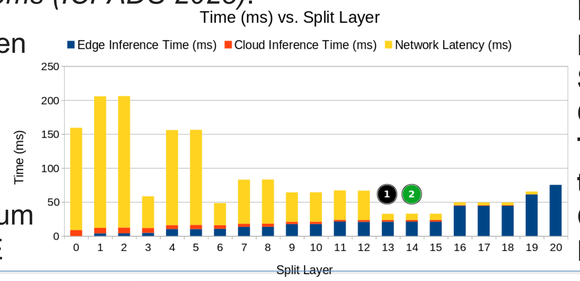


Niknami, N., Sawwan, A., & Wu, J. (2023, December). **SmartPipe: Intelligently Freezing Layers in Pipeline Parallelism for Distributed DNN Training**. In *Proc. of the IEEE International Conference on Parallel and Distributed Systems (ICPADS 2023)*.

M. Amin, H. Konjeti, J. Park, & Y. Wang. (2024-04-09). **An Enhanced Obstacle Detection in ADAS Applications by Integrating C-V2X with a Stereo Camera Vision System.** SAE WCX 2024, Detroit, MI, ISSN: 0148-7191.

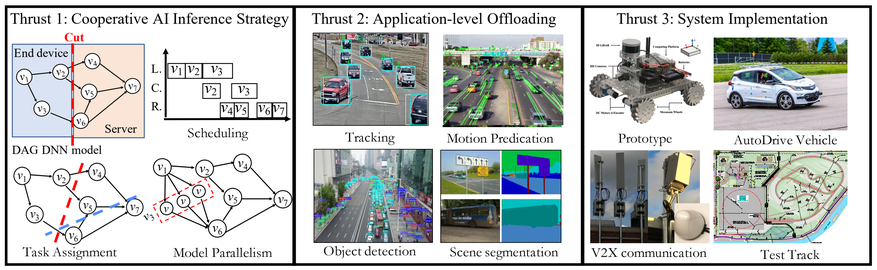


Purva Makarand Mhasakar, Kevin Doshi, Ning Wang, Shen Shyang Ho, and Haibin Ling, **Distributed Tracking and Verifying: A Real-Time and High-Accuracy Visual Tracking Edge Computing Framework for Internet of Things,** EdgeComm: The Fourth Workshop on Edge Computing and Communications (at ACM/IEEE Symposium on Edge Computing), December 9, 2023, Wilmington, DE



Nicholas Bovee, Stephen Piccolo, Suraj Bitla, Gopi Krishna Patapanchala and Shen-Shyang Ho, **SplitTracer: A Cooperative Inference Evaluation Toolkit for Computation Offloading on the Edge,** 8th IEEE International Conference on Fog and Edge Computing, Philadelphia, PA, May 6-9, 2024

Research Agenda:



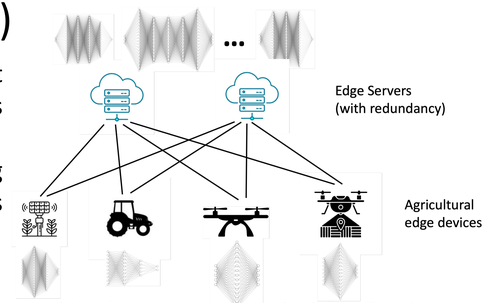
- **Thrust 1: Layer-wise Cooperative AI Inference Strategy Design.**
- **Thrust 2: Application-Level Offloading.**
- **Thrust 3: Vehicular Edge Computing Platform Development.**

Broader Impact (Education and Outreach):

- A new educational testbed integrated course with hands-on experiments to train next-generation intelligent transportation systems workforce.
- Undergraduate and graduate research programs: REU (Temple U), SURP (Rowan U), SAE International AutoDrive Challenge, etc.
- Integration of research into existing course curriculum (e.g., Machine Learning, Introduction to Artificial Intelligence for Autonomous Driving)

Broader Impact (Societal)

- Supporting AI application deployment in a variety of Cyber-Physical Systems (e.g., Agriculture technology).
- Eliminating the high-cost of installing high-end onboard computation units for future autonomous vehicles.



Shen-Shyang Ho, Paolo Rommel Sanchez, Nicholas Bovee, Suraj Bitla, Gopi Krishna Patapanchala and Stephen Piccolo, **Computation Offloading for Precision Agriculture using Cooperative Inference**, 8th IEEE International Conference on Fog and Edge Computing, Philadelphia, PA, May 6-9, 2024

Award ID#: 2128341, 2128378, 2128350, and 2128346