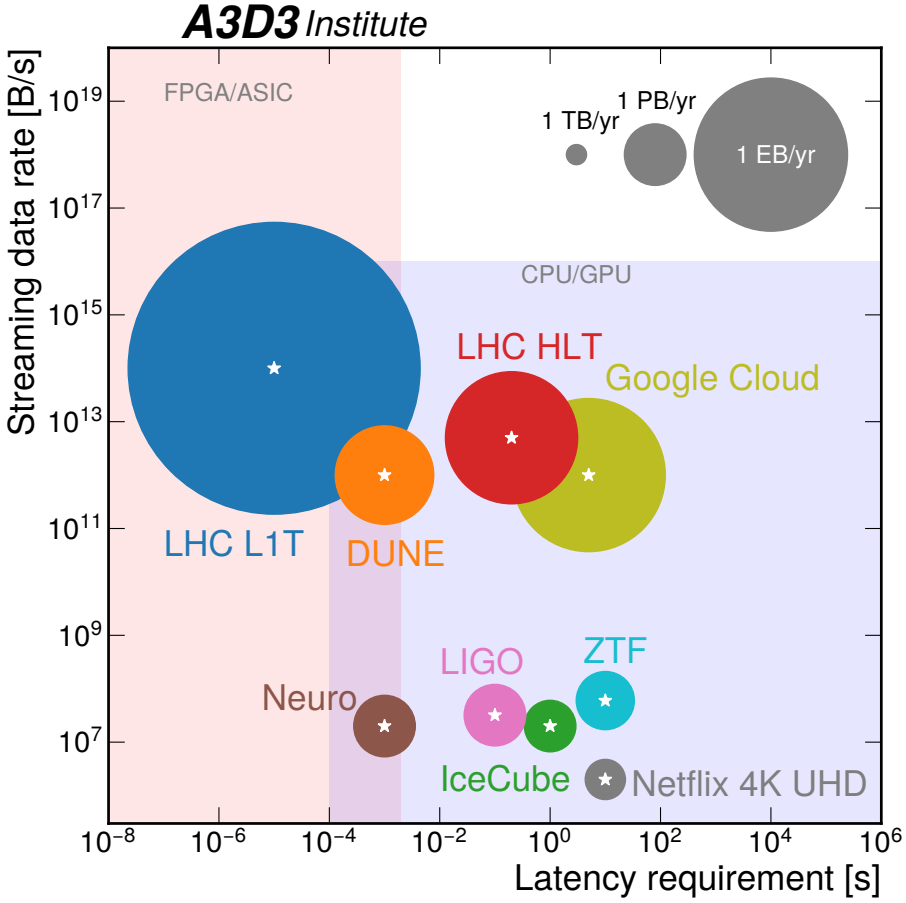
The Annual Report of Research Direction Algorithms and Hardware

Deming Chen, Song Han, **Pan Li**

10/07/2022

The Research Goal

- Large-scale scientific data introduce *new challenges*.
- Our goal: Prototype new *algorithms and hardware* for domain scientists to deal with such new challenges.



Main Teams and Members

- [UIUC] --- New Generation of Hardware Compilers and Design Automation

  Faculty: Deming Chen; Students: Hanchen Ye, Jialiang Liu

- [MIT] --- Algorithm-System-Hardware Co-Design for Efficient Point Cloud Processing

  Faculty: Song Han; Students: Zhijian Liu, Haotian Tang, Yujun Lin

- [Purdue] --- Interpretable and Generalizable Graph and Geometric Machine Learning

  Faculty: Pan Li; Students: Siqi Miao, Shikun Liu, Tianchun Li

Overall, 3 faculties, 8 students

Achievements

- [UIUC]

  Publication: 1 paper in HPCA'22; Several tutorial/workshop talks on ScaleHLS;
  Software products: ScaleHLS; Coordinating with HLS4ML.

- [MIT]

  Publication: 1 paper in MLsys'22; One paper (BVFusion) in submission
  Software products: TorchSparse 2.0; BVFsion

- [Purdue]

  Publications 1 paper in ICML'22; 1 paper in NeurIPS'21 workshop; 1 paper in VLDB'22

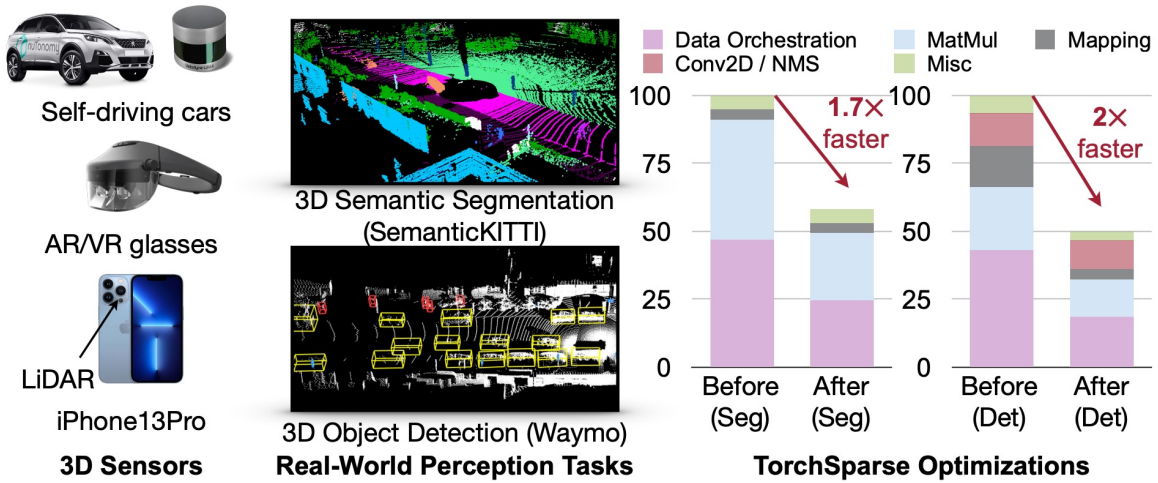
  Software products: GSAT; SUREL.

Overall, 6 papers in top venues, 5 github repos with total star # > 1.75k

Some Highlights of Projects

➢ [MIT] Torchsparse

How to pipeline 3D convolution computation on irregular 3D point cloud data?

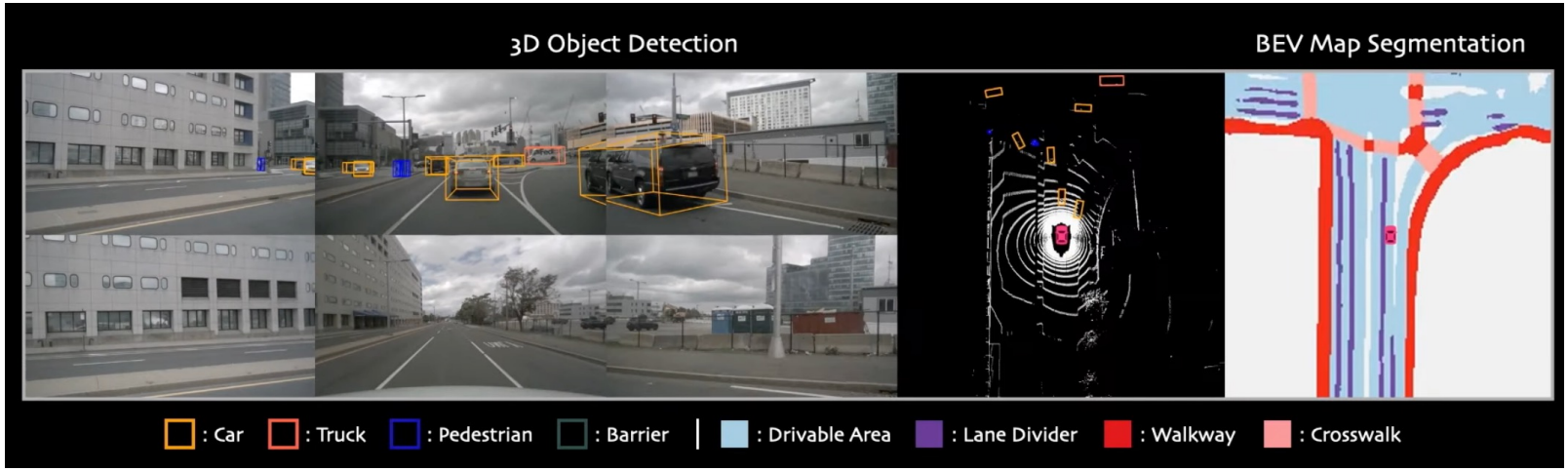


The model has been applied to HCAL and HGCAL and outperforms baselines largely.

Some Highlights of Projects

➢ [MIT] BVFushion

multi-task multi-sensor fusion: How to efficiently combine Vision + Lidar

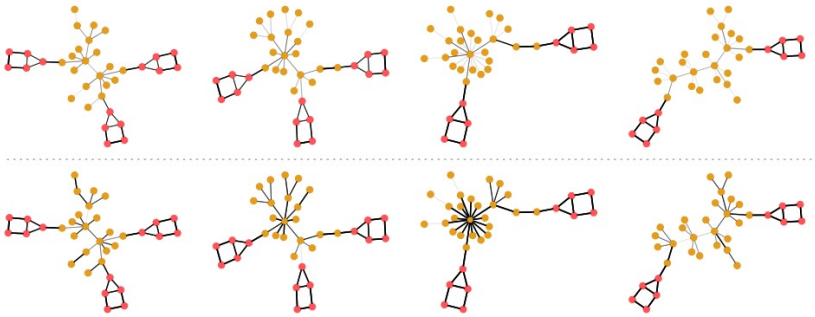


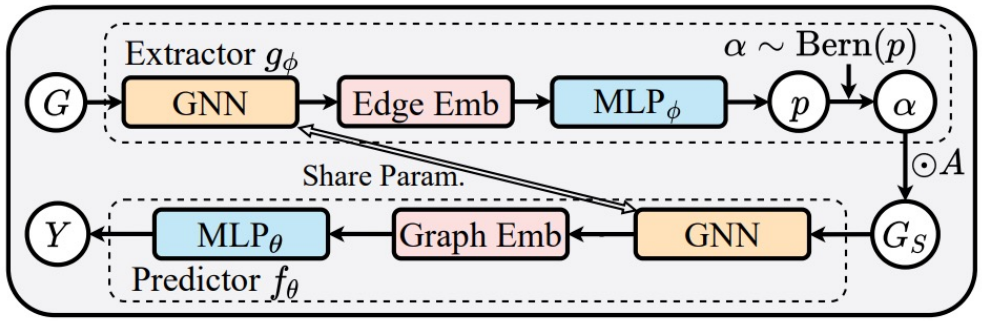
- Ranked 1st on nuScenes *3D object tracking* benchmark.
- Ranked 1st on Waymo *3D object detection* benchmark.

Some Highlights of Projects

➢ [Purdue] Graph Stochastic Attention

How to build inherently interpretable graph neural network models ?



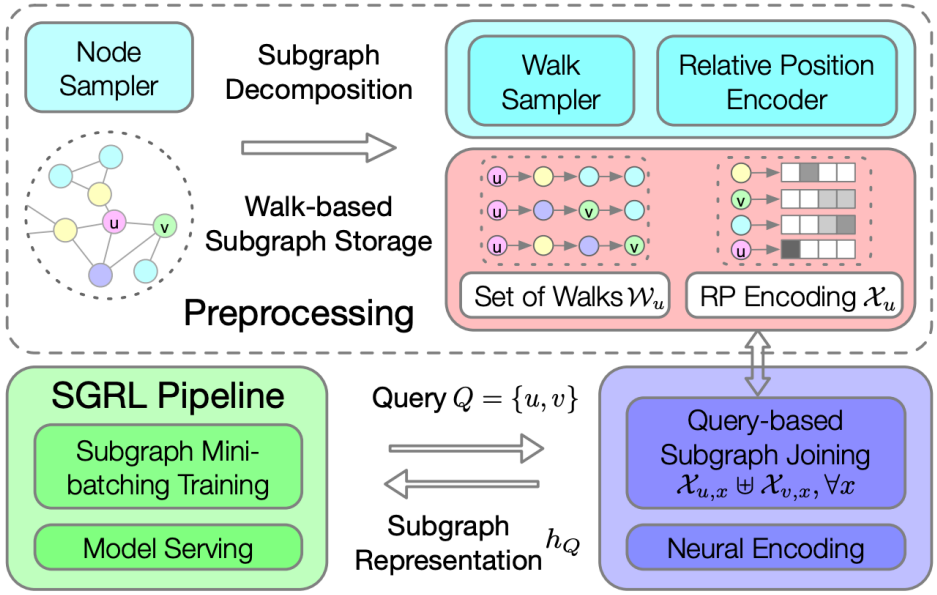


- Theoretically grounded by the principle of information bottleneck
- Outperform baselines in both interpretability and generalizability

Some Highlights of Projects

➢ [Purdue] Algorithm-System Co-design for Subgraph Representation Learning

How to build learn representations of subgraphs in large networks efficiently?

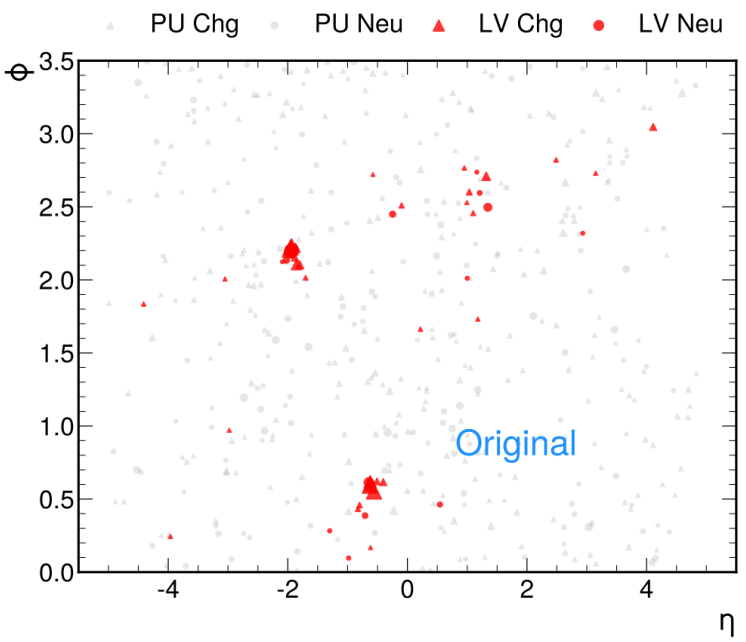


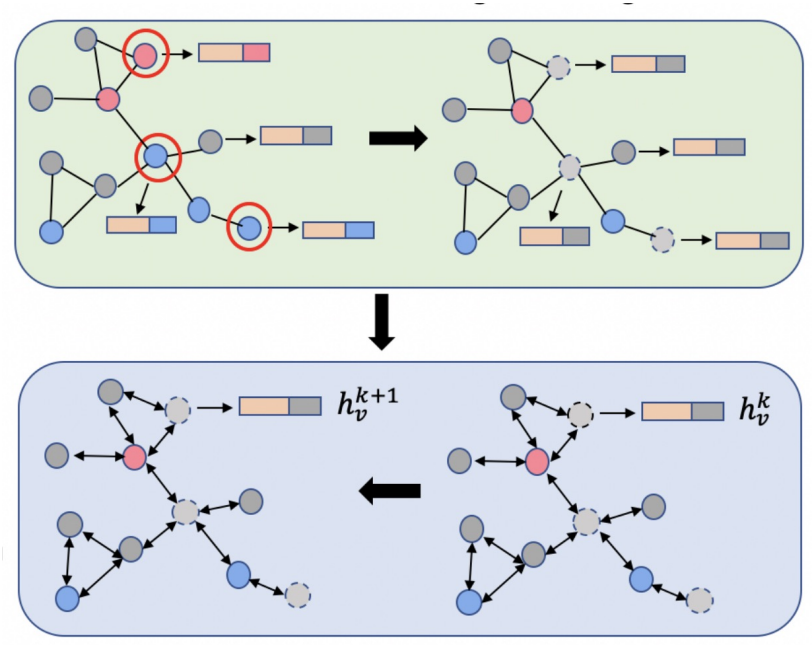
- Ranked 1st on link prediction on Open Graph Benchmark.

Some Highlights of Projects

- [Purdue] Semi-supervised Graph Neural Networks for Pileup Mitigation

How to address the problem where the labels of neutral particles are unavailable?





Future Plans

- [Torchsparse] (1) Optimize the efficient point cloud network SPVCNN++ for physical application (HCAL and compare with PFlow).
- [BEVFusion] (1) Evaluate BEVFusion on more 3D benchmarks and (2) deploy BEVFusion on NVIDIA Jetson AGX Orin 3.
- [GSAT] (1) Build the point-cloud interpretable model PSAT. (2) Build up a benchmark for point cloud data interpretation.
- [SUREL] (1) Accelerate SUREL by removing repetitive nodes in the sampled works. (2)Evaluate the model on new scientific applications, such as Brain Vessel Prediction
- [SSL-GNN] (1) Evaluation and finetune for real simulation data. (2) Build more generic domain adaptive graph neural networks.