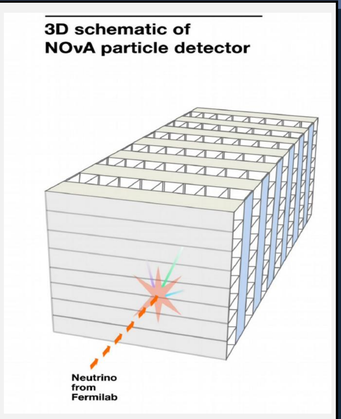
Prong Segmentation using Point Set Transformers in Multiple View Neutrino Detectors

Jiaxi Liu[1], Alejandro Yankelevich[1], Edgar E. Robles[2], Dikshant Sagar[2], Pierre Baldi[2], Jianming Bian[1], Wenjie Wu[1]
[1]Department of Physics and Astronomy, UC Irvine
[2]Department of Computer Science, UC Irvine

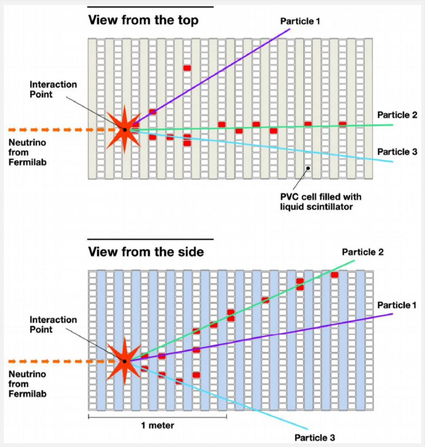
I. Background



- The **N**uMI **O**ff-Axis ν_e **A**ppearance Experiment (**NOvA**) is a long-baseline neutrino oscillation experiment detecting neutrinos from Fermilab's NuMI beam using liquid scintillator-filled PVC cells in planes of alternating orientation.
- Data: Two sparse 2D views (XZ top, YZ side) rather than native 3D.
- Task: Panoptic segmentation of particle 'prongs' (tracks or showers):
  - **Instance segmentation**: separate different prongs;
  - **Semantic segmentation**: classify particle types.

II. Motivation

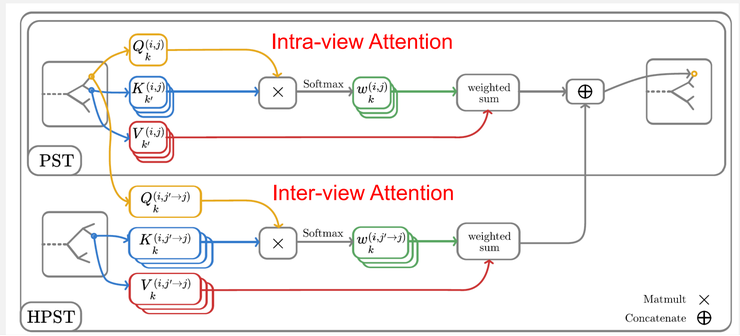
- ☐ Sparse detector images waste memory if using CNN.
- ☐ Two views have mismatched coordinates, preventing direct fusion.
- ☐ The new method must operate under limited resources.



Heterogeneous **P**oint **S**et **T**ransformer (**HPST**) was chosen for this task:
✓ Operates directly on sparse data by representing hits as sparse point clouds (coordinates + charge);
✓ Enables cross-view information flow with a heterogeneous attention mechanism;
✓ Good performance under limited computational resources by limiting attention to k-Nearest Neighbors (k-NN).

III. HPST Architecture

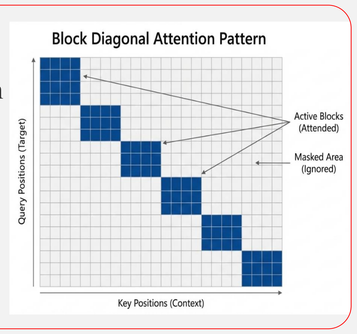
- Intra-view Attention: limited to k-NN, with Relative Position Embedding (RPE) to encode spatial information;
- Inter-view Attention: use a different pair of vectors, nearest neighbors calculated via shared Z coordinate.

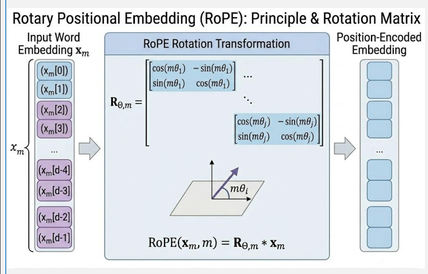


IV. Architecture Updates

FlashAttention Integration:

- Store data in block-diagonal attention pattern for efficient computation;
- Enables dense all-to-all attention; (from just using k-NN points)
- Significantly reduces training time & memory.



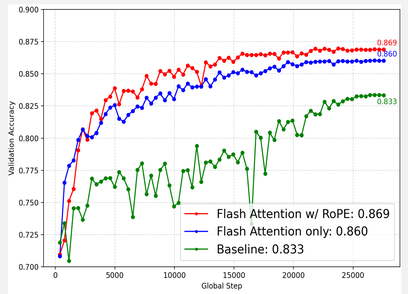


Rotary **P**ositional **E**mbedding **(RoPE)** replacing RPE:

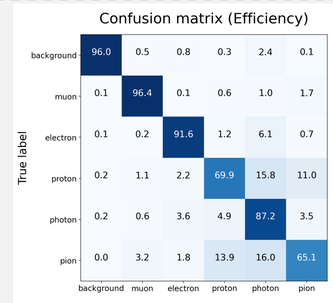
- Encode absolute position (m) as rotation, relative position $(m-n)$ as a 2×2 relative rotation matrix $R_{\theta(m-n)}$;
- Effectively preserve spatial relationships without extra computational cost.

V. Performance Benchmark

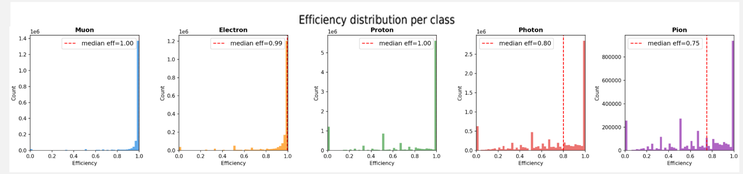
- Dataset: ~900k simulated NOvA events.
- HPST largely outperforms other models we used (R-CNN, GAT) while maintaining low computational resource consumption.
- Architecture updates (FlashAttention&RoPE) further improve the results.



Impact of architecture updates on validating performance



Semantic segmentation performance with updated architecture



Instance segmentation performance with updated architecture

VI. Application to LArTPC

- HPST generalizes to other multi-view detectors, such as LArTPC.
- Dataset: simulated 2m×7m×7m LArTPC with square 5mm pixel-based readout, ~200K events in total.
- 2D HPST surpasses other 2D models with fewer resources, while 3D PST has the best performance at the cost of increased memory usage.

Model	Memory (MiB)	Time (s)	Semantic Segmentation AUC	Instance Segmentation Accuracy
2D R-CNN	440.50 ± 51.04	1.575 ± 0.091	0.526	0.518
2D GAT	88.60 ± 7.56	0.230 ± 0.025	0.833	0.659
3D GAT	506.10 ± 30.13	0.722 ± 0.060	0.859	0.727
2D HPST (ours)	**99.10** ± 7.39	**0.354** ± 0.019	0.936	0.779
2D PST (ours)	138.10 ± 11.29	**0.254** ± 0.021	0.949	0.827
3D PST (ours)	**170.20** ± 9.65	0.140 ± 0.012	**0.982**	**0.889**

1. E. E. Robles, et al. Heterogeneous Point Set Transformers for Segmentation of Multiple View Particle Detectors. In ML4PS Workshop at NeurIPS, 2025.
2. E. E. Robles, et al. Particle hit clustering and identification using point set transformers in liquid argon time projection chambers. JINST, volume 20, P07030, 2025.
3. H. Zhao, L. Jiang, J. Jia, P. H. S. Torr, and V. Koltun. Point transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 16259–16268, 2021.
4. T. Dao, et al. Flashattention: Fast and memory-efficient exact attention with io-awareness. *Advances in Neural Information Processing Systems*, volume 35, pages 16344–16359, 2022.
5. J. Su, et al. Roformer: Enhanced transformer with rotary position embedding. Neurocomputing, volume 568, 127063, 2024.

Acknowledgement

This document was prepared by the NOvA Collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359.