Track Reconstruction in SPD Time Slices Using Graph Neural Networks with Contrastive Learning

A. Didorenko, N. Voytishin, S. Omelyanchuk*

*Joint Institute for Nuclear Research, Dubna, Moscow Region, Russia

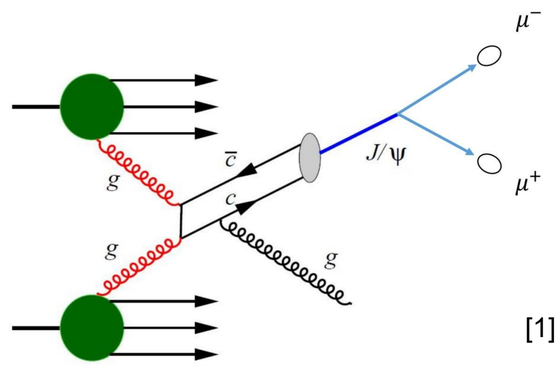
Abstract

The time slice decomposition in the SPD experiment at the NICA collider involves two stages: track reconstruction with noise removal followed by event separation, where high reconstruction efficiency is critical because each time slice contains 10–40 overlapping events. We propose a track reconstruction method based on graph neural networks with contrastive learning. A k-nearest-neighbor graph is built in hit coordinate space, and a neural network using EdgeConv layers trained with contrastive loss produces an embedding space where hits from the same track cluster compactly. Final clustering is performed with DBSCAN on simulated data from the official SPD detector simulation (SpdRoot). The method achieves high track reconstruction efficiency and is a promising candidate for the primary tracking algorithm in the first stage of time slice processing.

Dataset

Official SpdRoot simulation:

- Process: $J/\psi \to \mu\mu$
- Detector: Straw Tracker
- Events: 70,000
- Split: 80% train / 10% val / 10% test
- Time slices: constructed from separate events

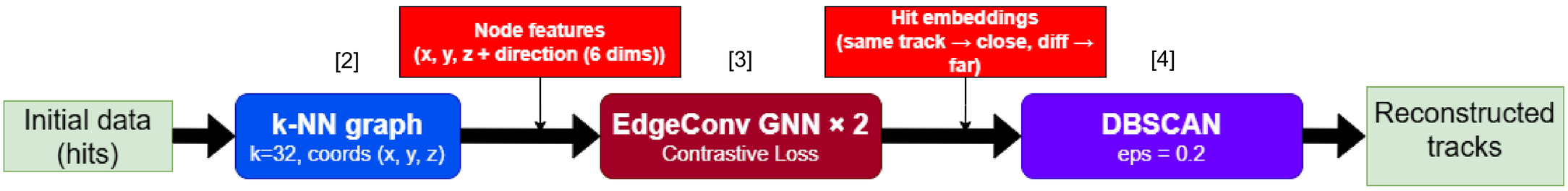


Software & Hardware

[1]

- Data simulation: SpdRoot (NICA NCX cluster)
- Processing: HybriLIT JINR, GPU Tesla V100 32GB

Pipeline

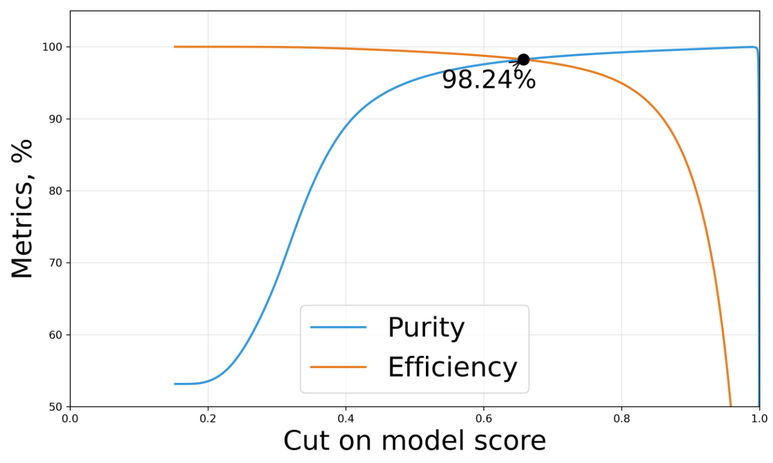


Results

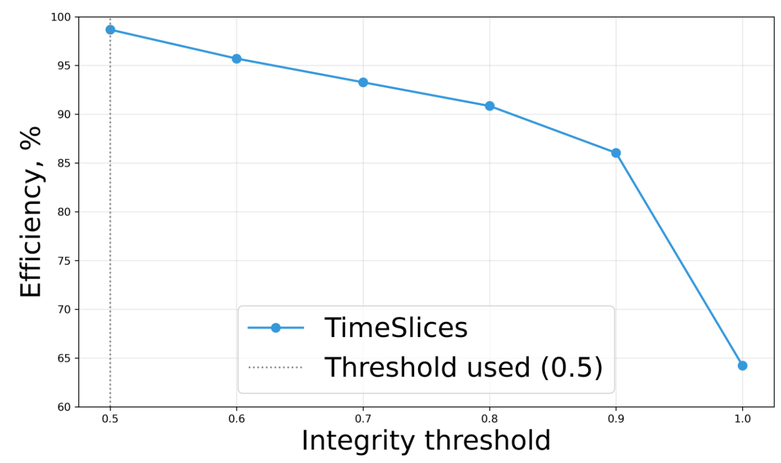
Track reconstruction efficiency is a standard HEP metric measuring how well an algorithm reconstructs simulated tracks. It is defined as:

$$Eff = \frac{N_{reco}}{N_{sim}} * 100\%,$$

where N_{reco} is the number of reconstructed tracks and N_{sim} is the number of simulated tracks. Our pipeline achieves an efficiency of 98.67% with an average processing time of 0.65 s per time slice.

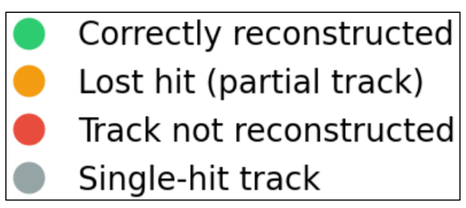


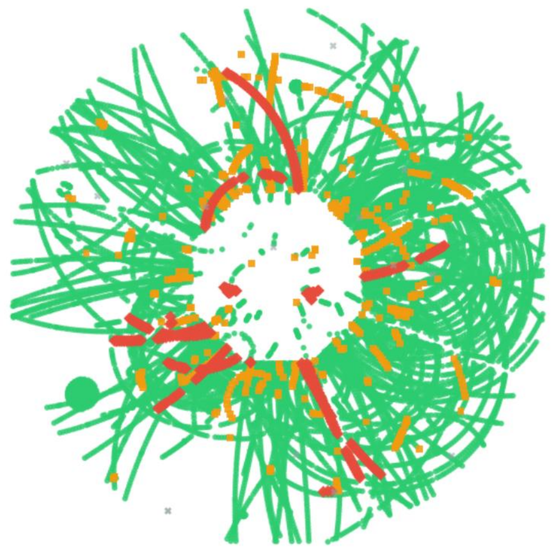
Purity-Efficiency



Efficiency vs Integrity threshold

Best reconstructed time slice





Worst reconstructed time slice

Conclusions

A track reconstruction pipeline combining GNN with contrastive loss and a DBSCAN clustering has been developed. The algorithm demonstrates excellent performance, achieving an efficiency of 98.67% with an average processing time of 0.65 s per time slice. The pipeline has been validated on the official SPD simulation (SpdRoot) using $J/\psi \to \mu\mu$ events, proving its applicability to realistic detector data.
Future plans include retraining the model directly on time slice data (currently trained on single events) and testing on more complex physics processes.

References

1. SPD Collaboration (V. Abazov et al.) – Technical Design Report of the Spin Physics Detector at NICA – arXiv – 2024 – 1-362, DOI: https://doi.org/10.48550/arXiv.2404.08317
2. Jon Louis Bentley. 1975. Multidimensional binary search trees used for associative searching. Commun. ACM 18, 9 (Sept. 1975), 509–517. https://doi.org/10.1145/361002.361007
3. Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, J. M. Solomon. Dynamic Graph CNN for Learning on Point Clouds. ACM Trans. Graph., Vol. 38, No. 4, Article 146, 2019. DOI: 10.1145/3326362
4. Ester, Martin et al. “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise.” Knowledge Discovery and Data Mining (1996).

MESHCHERYAKOV LABORATORY OF INFORMATION TECHNOLOGIES