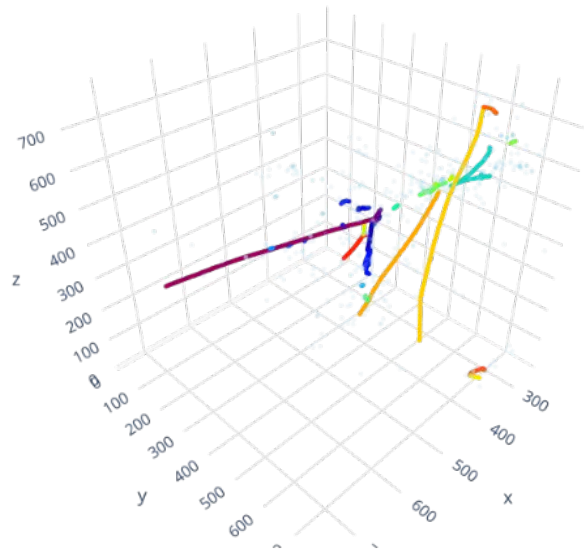
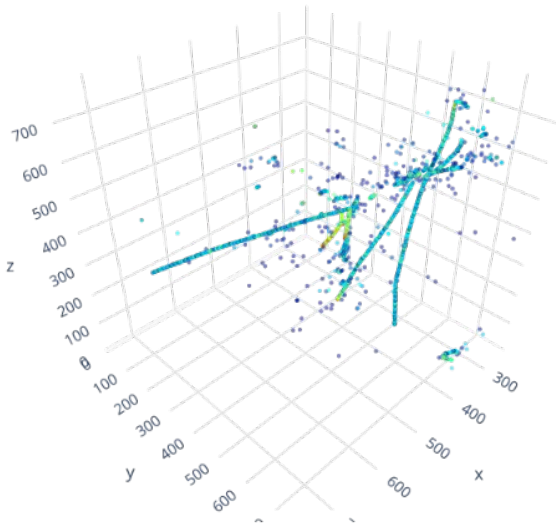


Particle Clustering in LArTPCs with Embedding Learning Convolutional Neural Networks

SLAC National Accelerator Laboratory / Stanford
Dae Heun Koh* / Kazuhiro Terao

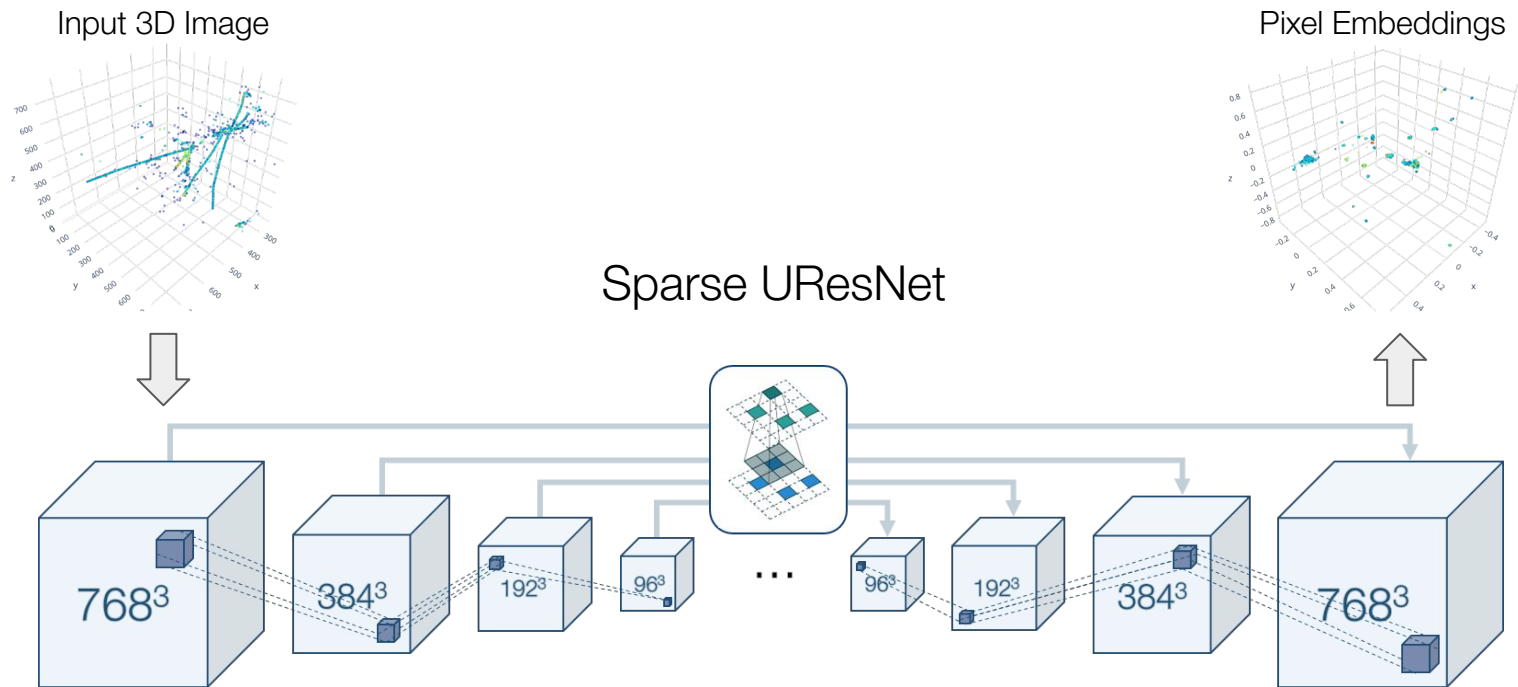
Overview

Instance Segmentation: identify each instance within an image at the pixel level



Overview

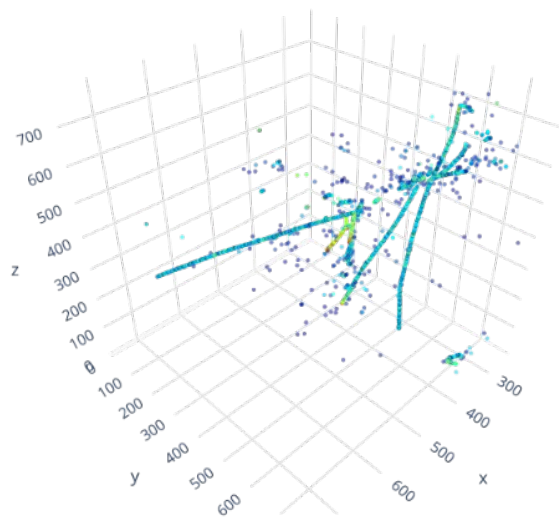
Idea: Use Convolutional Neural Network (CNN) as coordinate transform



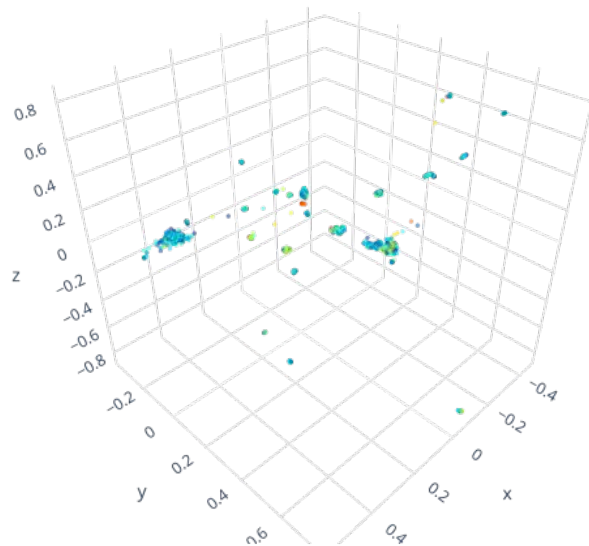
Overview

Idea: Pixel embeddings that belong to the same particle are embedded closer to each other.

Input 3D Image



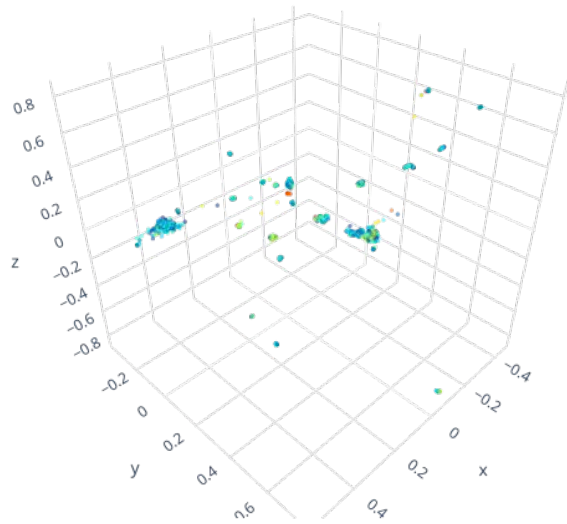
Pixel Embeddings



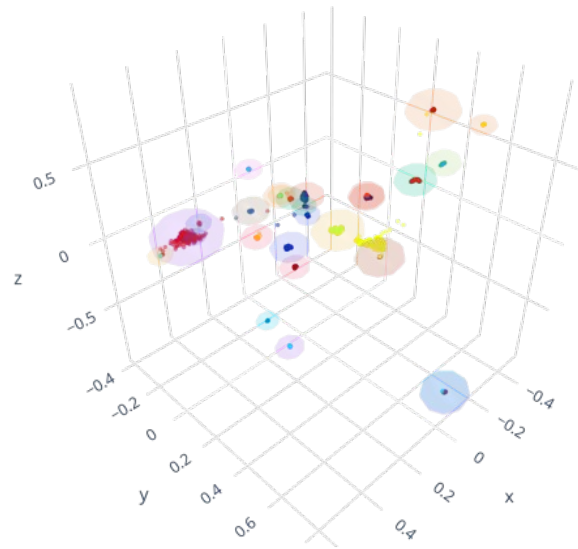
Overview

Idea: Once the embeddings are well-separated, we assign particle instance labels sequentially per cluster.

Pixel Embeddings



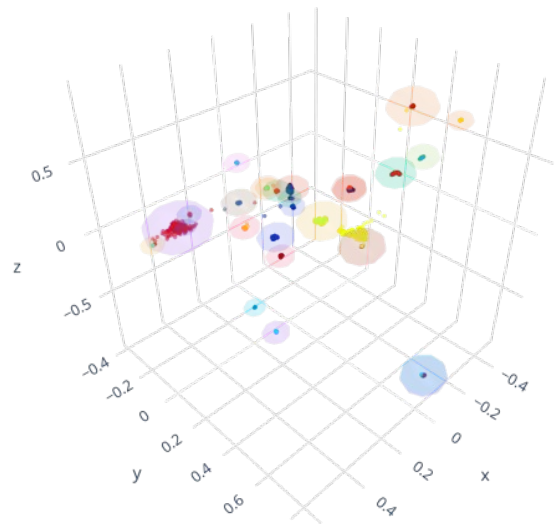
Post Processing



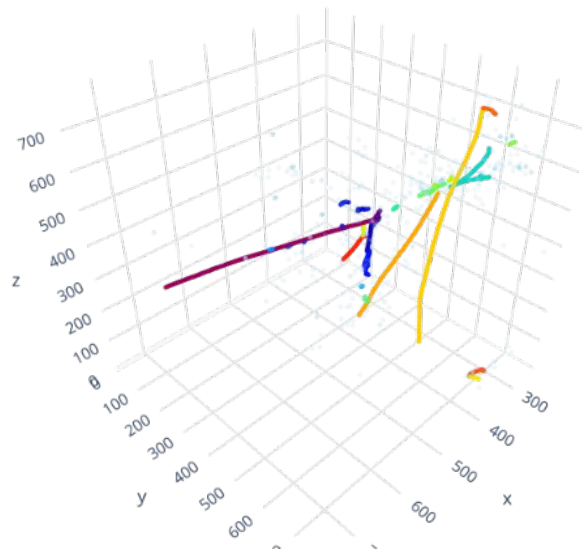
Overview

Idea: Since each pixel in embedding space is in one-to-one correspondence with the original image pixels, we translate the labels to the original image to obtain instance labels.

Pixel Embeddings

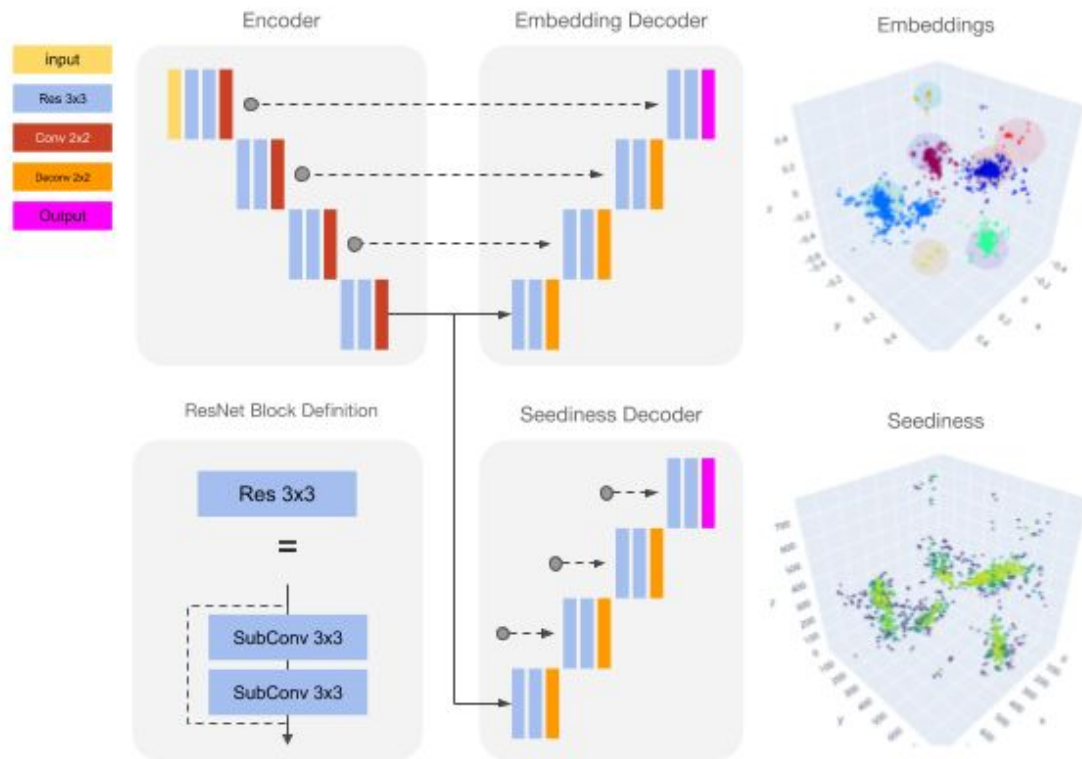


Instance labels



Network Architecture

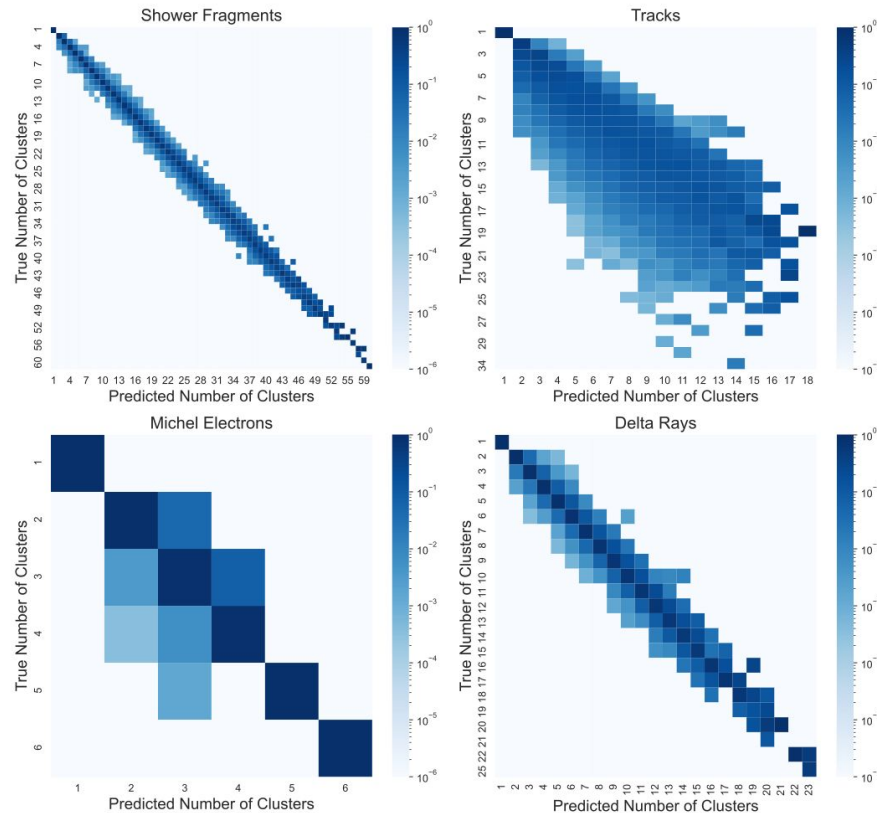
- ❑ **UResNet:** UNet with residual connections, implemented with sparse convolutional layers.
- ❑ **Embedding:** generate pixel embeddings from input image pixels.
- ❑ **Seediness:** score measuring a pixel's proximity to the instance centroid in embedding space.



Results

- Train and evaluate network on Monte-Carlo generated open dataset (PILArNet: <https://arxiv.org/abs/2006.01993>)
- 80k Training set, ~20k validation set.

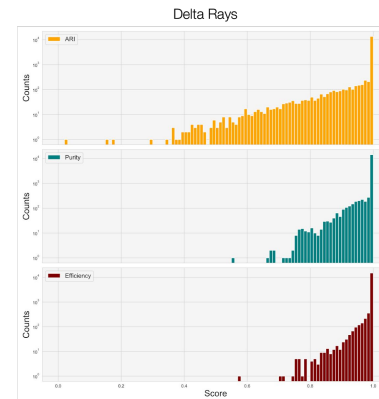
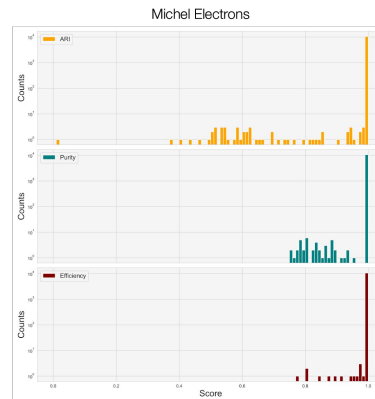
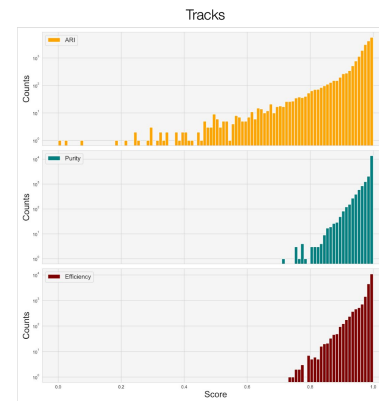
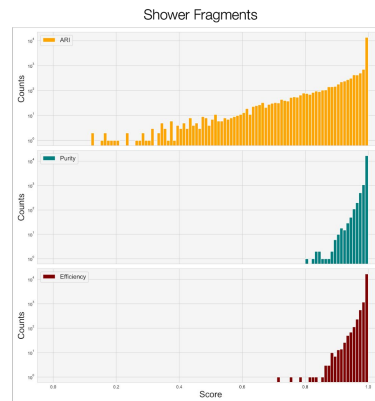
	Average # per Event	Voxel Counts	Percentage of Particle Instances
Shower Fragments	15 +- 9	21.4M	57%
Tracks	6 +- 3	47.7M	25%
Michel	1 +- 1	0.793M	3%
Delta	5 +- 3	1.11M	15%



Results

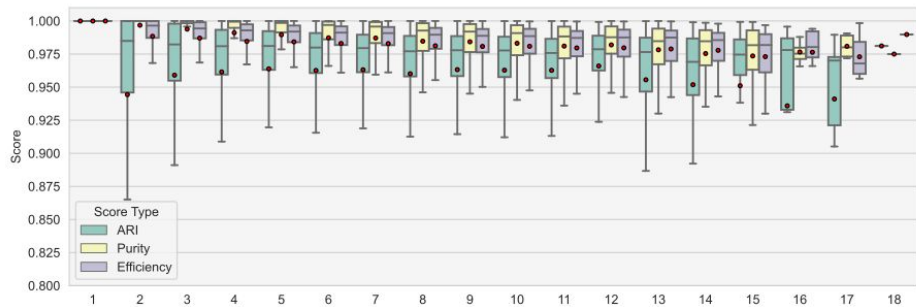
- ❑ Use ground truth semantic labels and cluster within each semantic class.
- ❑ Mean Purity and Efficiency is above **99%**
- ❑ Mean ARI = **0.973**

	Shower Fragments	Tracks	Michel	Delta
ARI	0.968	0.961	0.998	0.978
Purity	0.997	0.988	0.999	0.992
Efficiency	0.996	0.983	1.000	0.998
SBD	0.978	0.717	0.998	0.982

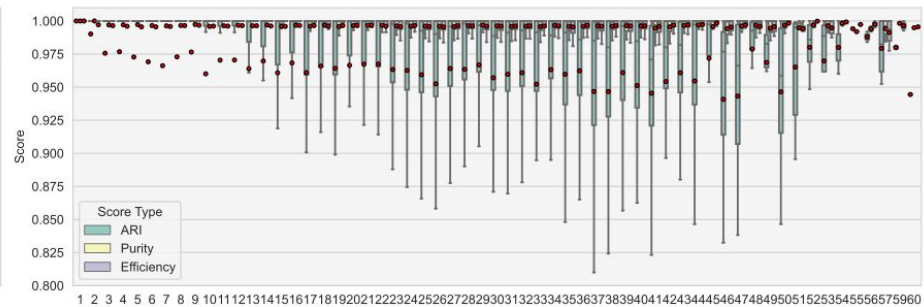


Results

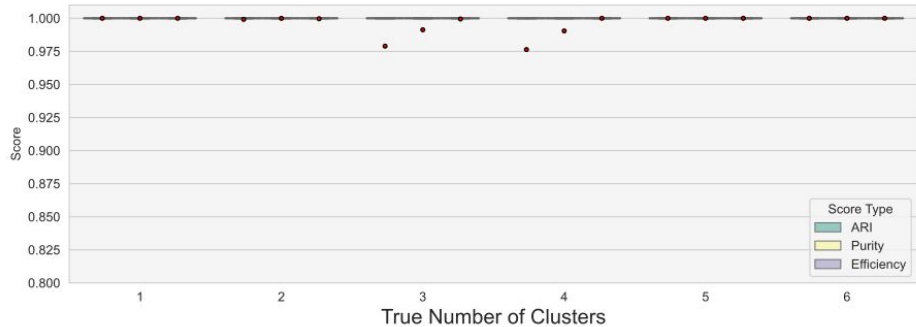
Track (HIP + MIP)



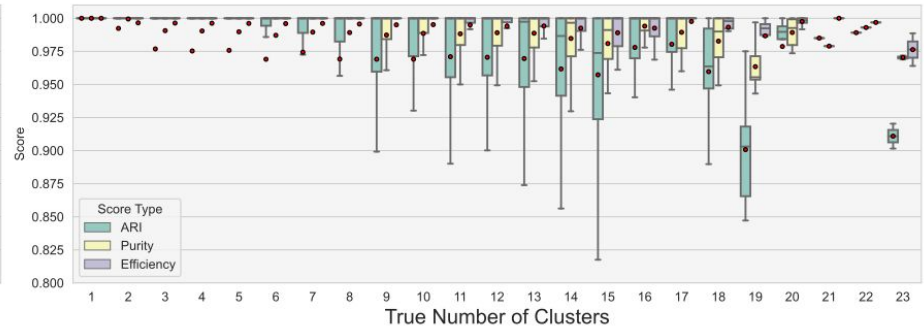
Shower Fragments



Michel Electrons

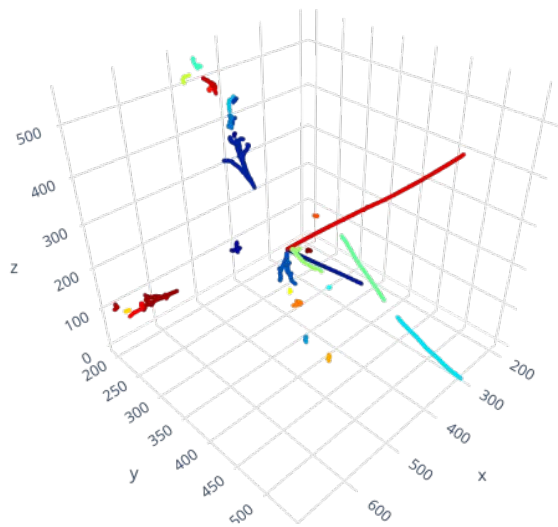


Delta Rays

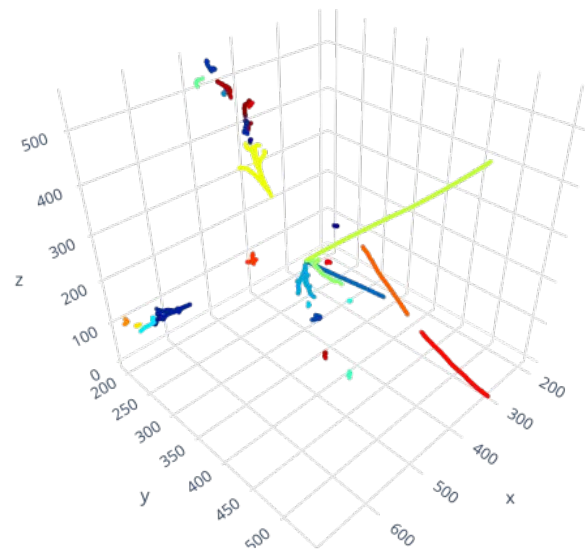


Results

Prediction

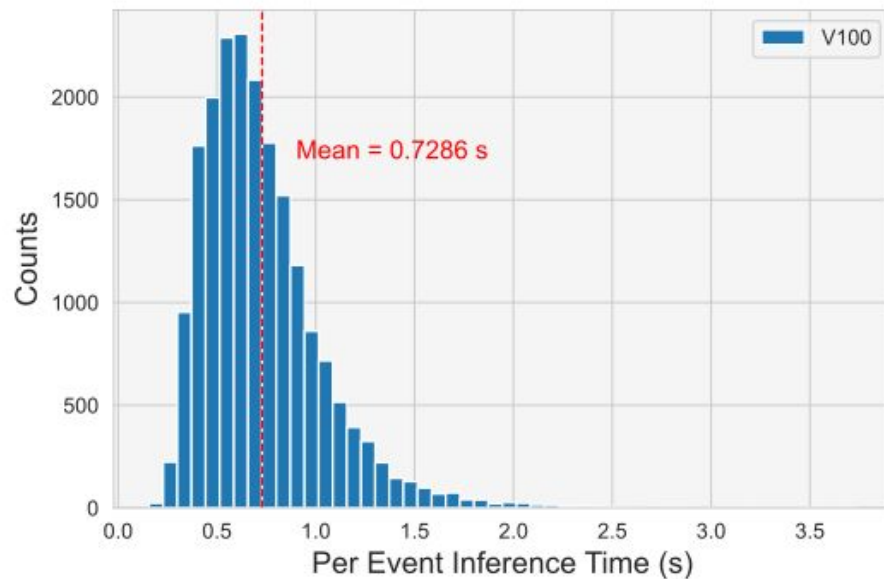


Truth



Results

- ❑ Training done in NVIDIA V100 with batch size 32
- ❑ Average train time per iteration ~ 10s
- ❑ Memory usage during training ~ 10GB
- ❑ Average time for label generation per image: ~0.7s



Future Directions

- ❑ Improvements for track clustering
- ❑ Integration with ghost removal and the whole ML reconstruction chain being developed in SLAC.
- ❑ Benchmark study with other ML based methods, namely Mask-RCNN.