DeepHalo Extension: Instance Segmentation of Dark Matter Halos

Rishav Das and Ipsita Rout
National Institute of Science Education and Research

Abstract

Our goal is to use GNN to instance segment 3D point cloud data. For that purpose we are using dark matter only data from TNGIllustris, the run is TNG50-4-Dark catalogue. Our goals up until this point were to become acquainted with the data schema, perform exploratory data analysis, and gain knowledge of GNN. All of that has been accomplished. The point cloud data has been shown using color-coded subhalos. We have picked halo number 30 for that reason. We intend to use octrees in the future to partition the point cloud data spatially so that gnn can be applied to it.

1 Analysis of Related Papers

1.1 PointGroup: Dual-Set Point Grouping for 3D Instance SegmentationJiang et al. 2020

https://doi.org/10.48550/arXiv.2004.01658

With regard to better grouping the points, PointGroup is a new end-to-end bottom-up architecture that investigates the empty space between objects. They use a two-branch network to extract point features and predict semantic labels. Both the original and offset shifted point coordinates are utilised. They ran it on datasets ScanNet v2 and S3DIS and have the highest performance rate of 63.6 % and 64.0 %.

1.2 Divide and Conquer: 3D Point Cloud Instance Segmentation With Point-Wise BinarizationZhao et al. 2023

https://doi.org/10.48550/arXiv.2207.11209

The previous paper used a distance clustering framework. However, it has difficulties when attempting to partition points with identical semantic labels, particularly when adjacent points are clinging to one another. This paper claims to solve that problem by utilising a framework where they binarise each point and assign them different clusters. They also ran it on datasets ScanNet v2 and S3DIS.

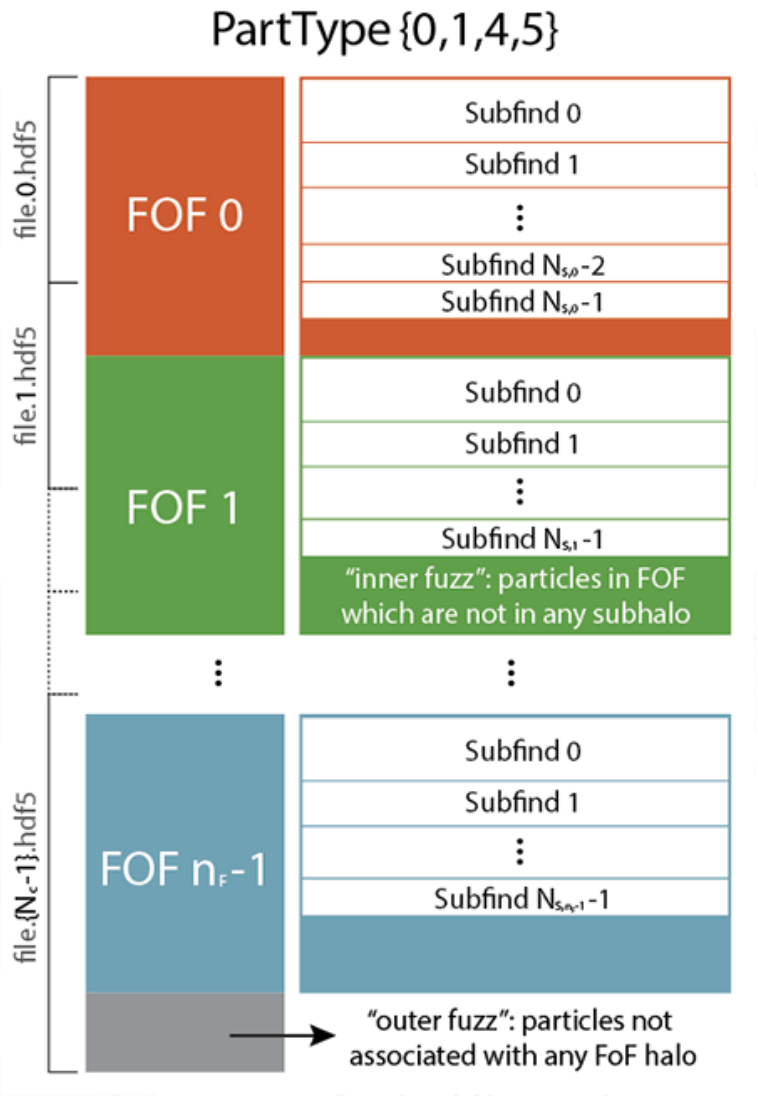
1.3 Point-GNN: Graph Neural Network for 3D Object Detection in a Point CloudShi, Ragunathan, and Rajkumar 2020

https://doi.org/10.48550/arXiv.2003.01251

Now, this paper uses GNN for object detection on LiDAR data. This can be used for Instance Segmentation.

1.4 Exploaratory Data Analysis

Each run has a 100 snapshots, from t = 0 and z = 0.



In each snapshot, data is stored as such: each FOF here refers to a halo. The subhalos are stored in descending order of size, the first subhalo being the central subhalo, with the largest number of particles. Fuzz, or particles that do not belong to any halos are stored at the end.

There are various particle types. Since we are using only dark matter files, we are only concerned with PartType 1.

We visualised the data, taling HaloID = 30 of TNG50-4-Dark and plotting it with it's subhalos colour-coded. This is what we should be expecting as a result from our machine learning model.

Here, we can see that the class-imbalanced. Also, there seems to be a bias here. Most of the subhalos are concentrated towards one side–the most probable explanation for this could be that there is a gravitational sink at that side.

Where the central subhalo (i.e. the largest subhalo of the halo is considered part of the background and hence, colour-coded as such).

1.5 Future Plans

Instance segmentation has been successfully implemented for 2D data. Converting point-cloud data into graph would help us implement it on 3D point cloud data.

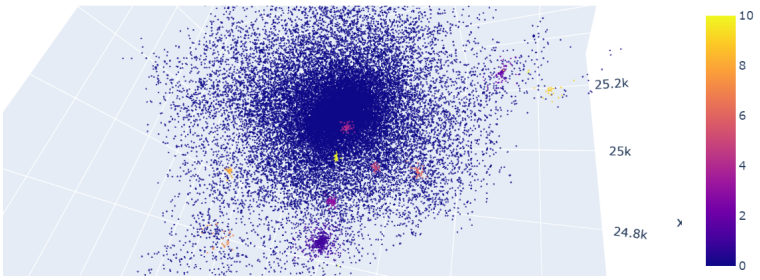


Figure 1:

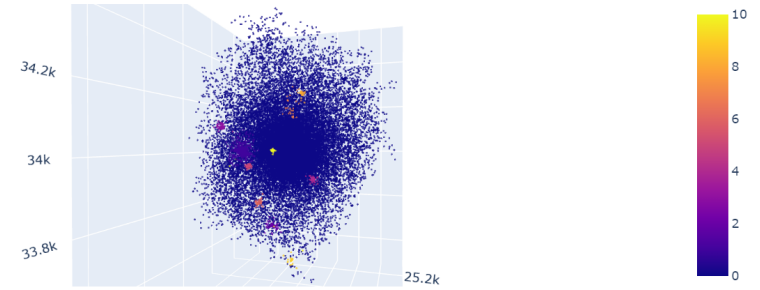


Figure 2:

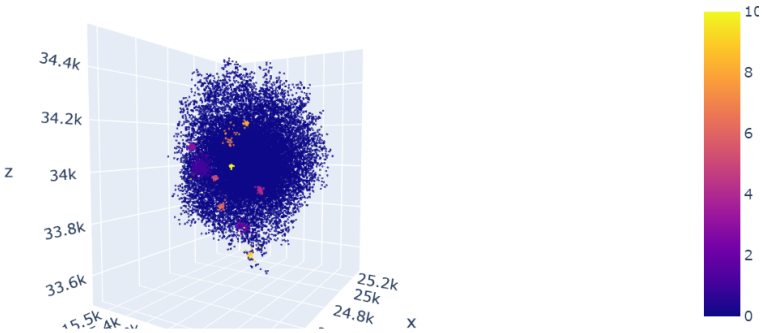


Figure 3:

We intend to use Octrees to convert the point cloud data into graph structure, following which we can run GNN on it.

References

Jiang, Li et al. (2020). *Pointgroup: Dual-set point grouping for 3D instance segmentation*. URL: https://arxiv.org/abs/2004.01658.

Shi, Weijing, Ragunathan, and Rajkumar (2020). *Point-GNN: Graph neural network for 3D object detection in a point cloud*. URL: https://arxiv.org/abs/2003.01251.

Zhao, Weiguang et al. (2023). *Divide and conquer: 3D point cloud instance segmentation with point-wise binarization*. URL: https://arxiv.org/abs/2207.11209.