111 學年度大學部專題競賽

國立清華大學資訊工程學系
Department of Computer Science, National Tsing Hua University

3D Instance Segmentation on Synthetic and Real Aerial Photogrammetry 3D Point Cloud Dataset

Megan Kuo（郭玟均）, *Louie Yen*(顏訓哲）

Introduction

Dataset

Specifically, we utilize the STPLS3D benchmark because it is used to provide synthetic urban scale point clouds with high-quality instance annotations. This dataset has more than 17 km of landscapes, up to 18 fine-grained semantic categories, and 15 instance categories provided, which is a very large dataset and might cost a lot of resources to do the computations.

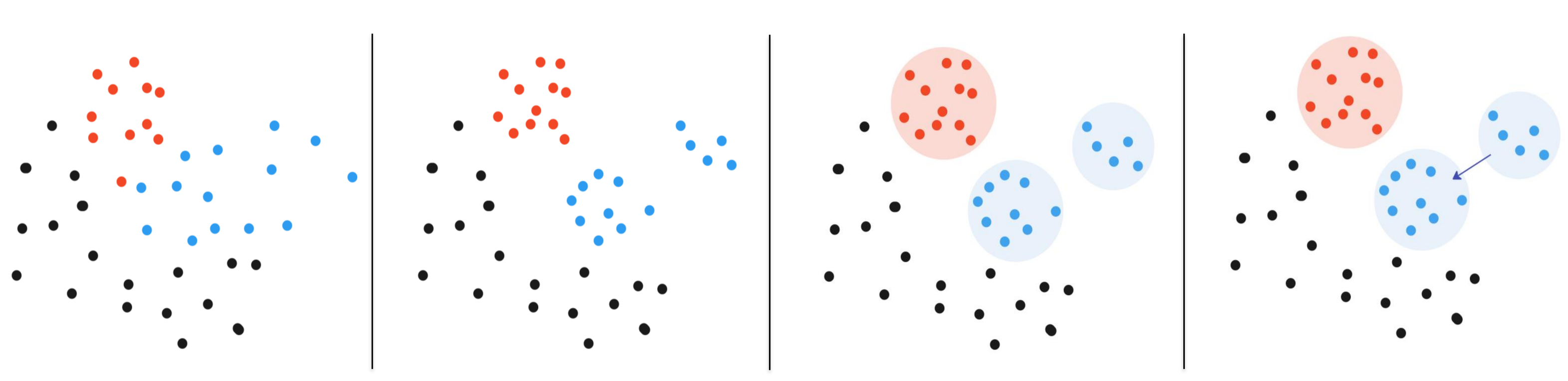
Research

Our research aims to explore how to achieve instance segmentation of 3D point clouds. While the current segmentation models have the defect of transferring the semantic prediction error into instance segmentation, which causes biased prediction, we have tried to explore ways to enhance the model's performance by focusing the model's attention more on the segmentation task of instance rather than semantic.

Material

Backbone: HAIS

We chose the SOTA model, the hierarchical aggregation model as the baseline of the STPLS3D dataset.



Process of Primary instance absorbing fragments.

Method

Sampling

We plan on dividing sampling units into different layers according to a certain rotation angle, attempting to obtain sample structures that are similar to that of the population distribution.

Dropout

For handling mispredicted weights, we wanted to find a way to provide our training model with a different view of configured layers, without adding the additional computational expense of the model.

Noise

We randomly picked the x, y, and z coordinates to add noise where random spots are replaced by extremely dark and bright values.

Attention

Observing that the false positive result produced from the semantic segmentation significantly affects the performance of the instance segmentation tasks, we plan to ameliorate this effect by extracting improbable information so that the model can make a more accurate prediction.

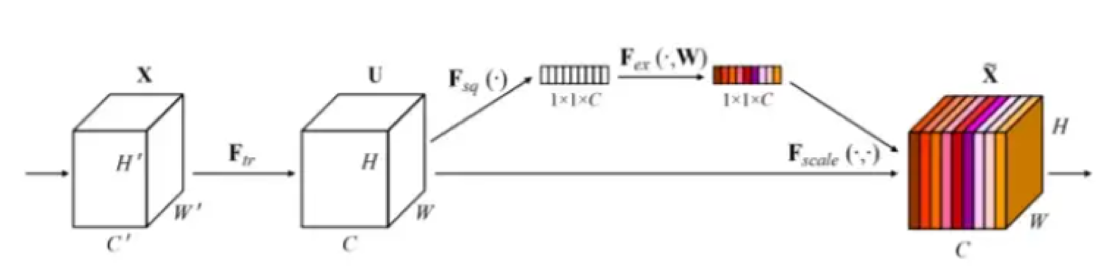
Simulation and tuning

Augmentations

- **Sampling**
  Using a stratified random sampling method, we divided sampling units into different layers according to a 30-degree rotation angle, and then randomly selected samples from different layers.
- **Dropout**
  Instead of removing single nodes or points, we randomly removed a whole 3D feature map; while with the probability sampled from the Bernoulli distribution, every channel has the chance to be zeroed out during the feed-forward process.
- **Noise**
  Random replace original points with black and white points.

Attention Layers

We introduced adaptive average pooling on channel*heights*weights to obtain a feature map of size 1*1*channel, then used the fully connected layers to perform a nonlinear transformation on the result.



Experiment Result

	Baseline AP	Data Augmentation AP	Attention AP
Building	0.533	0.524	0.688
Low Veg	0.122	0.092	0.236
Med Veg	0.112	0.097	0.227
High Veg	0.186	0.190	0.242
Vehicle	0.674	0.532	0.792
Truck	0.335	0.280	0.630
Aircraft	0.190	0.241	0.286
Military	0.063	0.114	0.471
Bike	0.019	0.030	0.141
Motorcycle	0.445	0.351	0.531
LightPole	0.391	0.198	0.418
Street Sign	0.052	0.049	0.110
Clutter	0.100	0.158	0.368
Fence	0.135	0.109	0.185
Average	**0.24**	**0.212**	**0.380**

Conclusion

Upon doing the experiment, we encountered a lot of technical problems, and among these problems, it is creating a virtual environment for the HAIS model that took the most time.

According to the previous experiment we conducted, adding new blocks into the model is more effective in the performance of the model than tuning parameters, but the cost is that it must spend more time on computing additional layers.

We are grateful to have this opportunity to present our project, thanks for the professor's feedback and the senior's help.