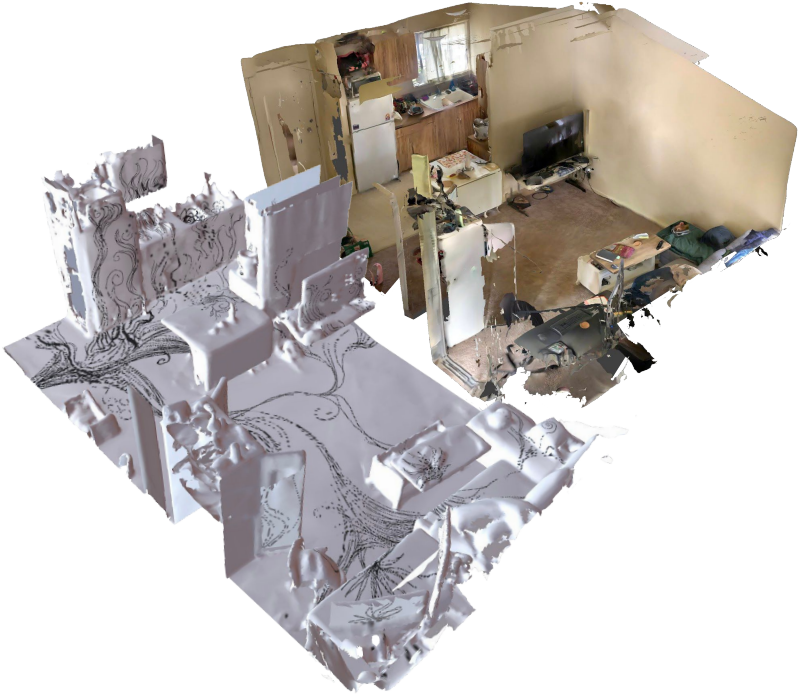
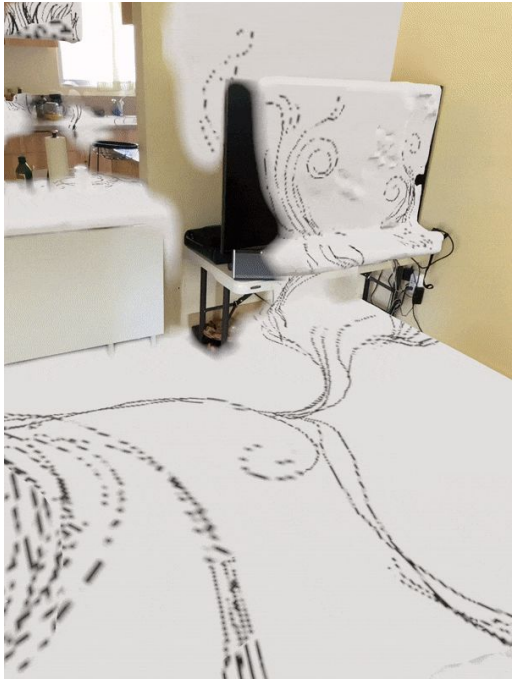
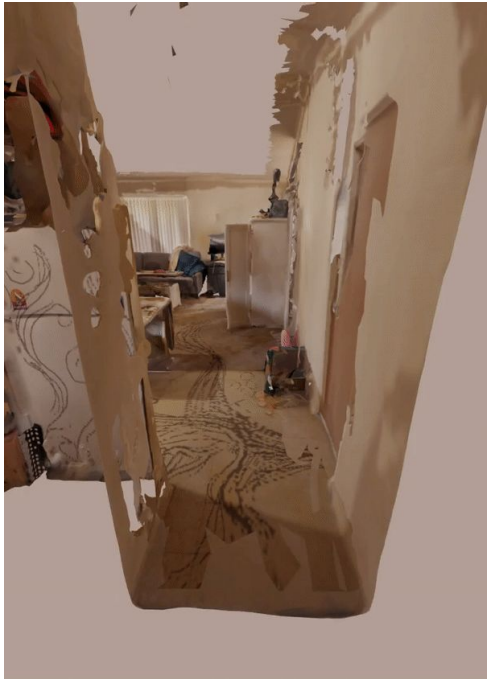
PointCloud segmentation a HCI / Media Arts perspective...

Slides: Ana M. Cárdenas

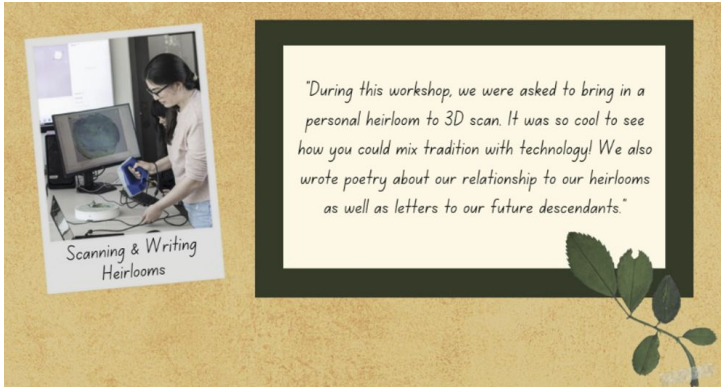
HCI for creativity support in AR







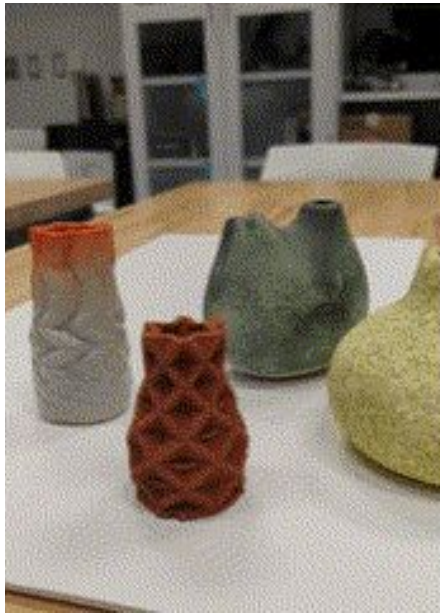
How can people easily create and edit 3D assets for immersive content?

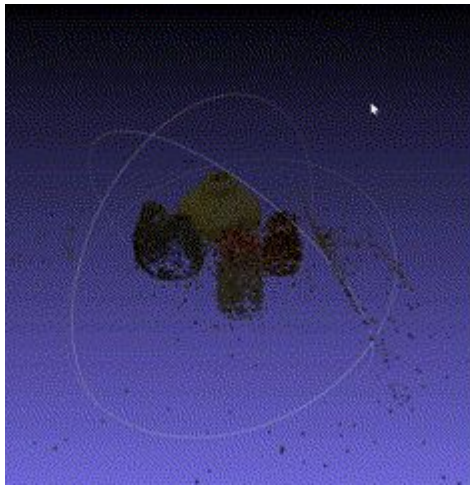


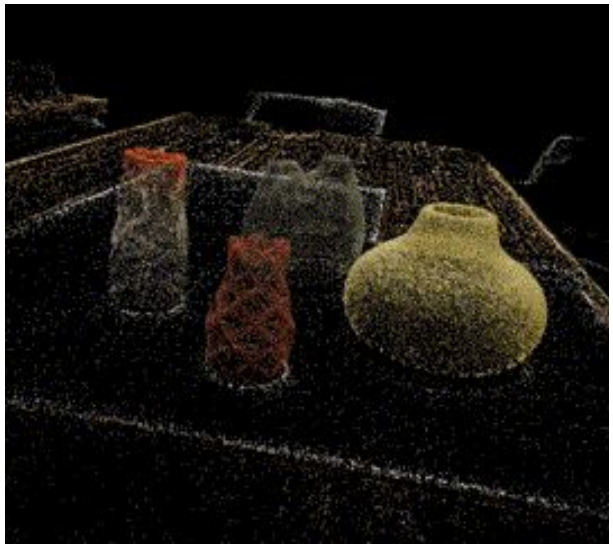
Why point clouds?

SFM (structure from motion) using ColMAP

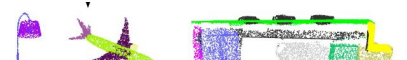
Gaussian Splatting

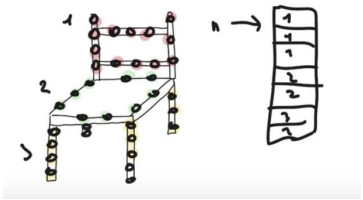






Point Cloud Segmentation



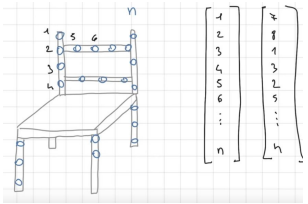


Segmentation

input: the input can be a single object for part region segmentation, or a sub-volume from a 3D scene for object region segmentation.
output: n × m scores for each of the n points and each of the m semantic sub- categories.

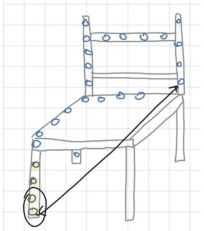
Unordered

Network should be Invariant to permutations of the input set of points



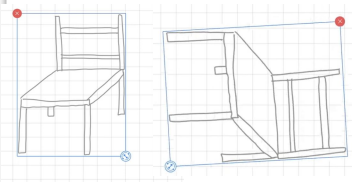
Interaction among points

Network should capture local structures from nearby points



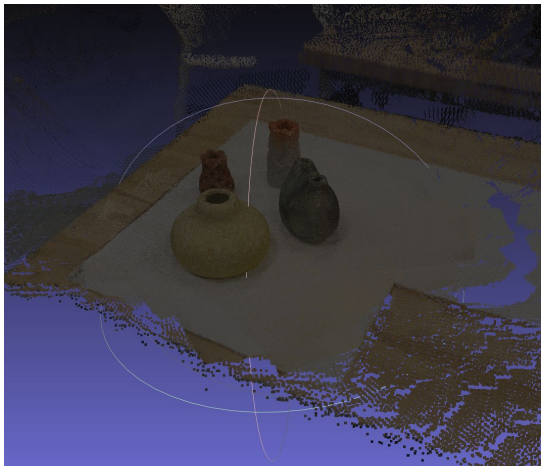
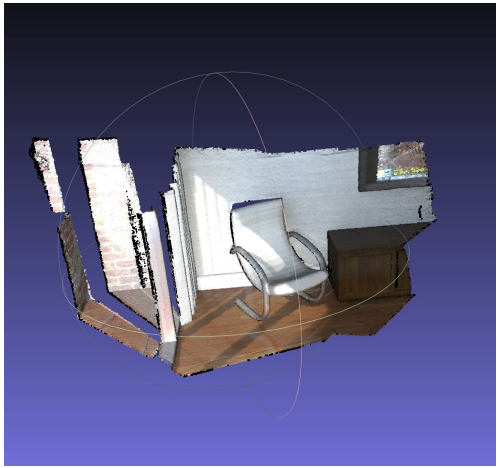
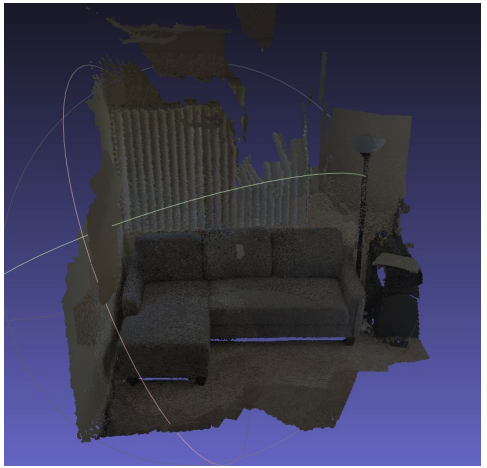
Invariance under transformations

Network should be invariant to rotating and translating set of points all together

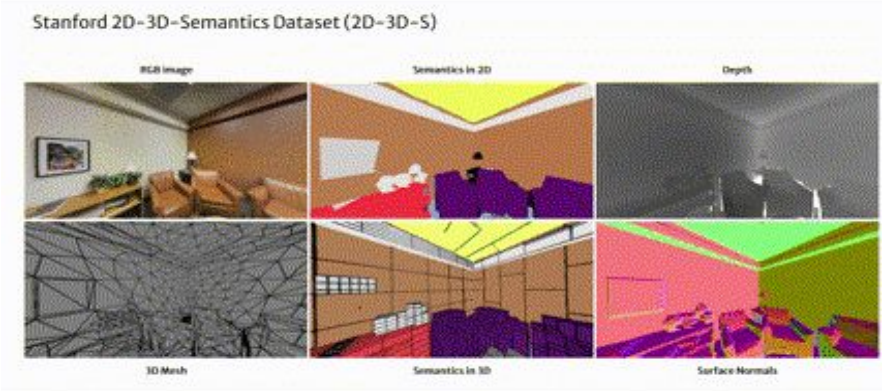


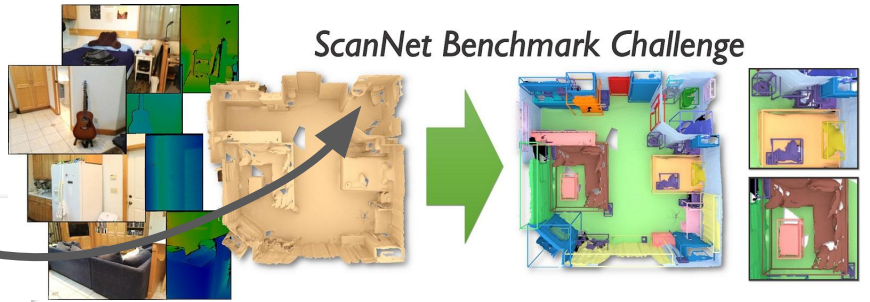
Point Cloud Segmentation could be useful to select objects from captured scenes or to create interactive 3D scenes...

Can I use pre-trained models with my own pointclouds?

Because of current datasets, most things we can segment with existing models are labs and offices 😞 (or streets for self-driving cars)





Fun Fact! Ellings and HFH are in this dataset

Focus on models trained with indoors dataset… ScanNet is impossible to download so I used S3DIS

https://github.com/isl-org/Open3D-ML

Semantic Segmentation on S3DIS

https://github.com/Pointcept/Pointcept

POINTCEPT
Point Cloud Perception Codebase

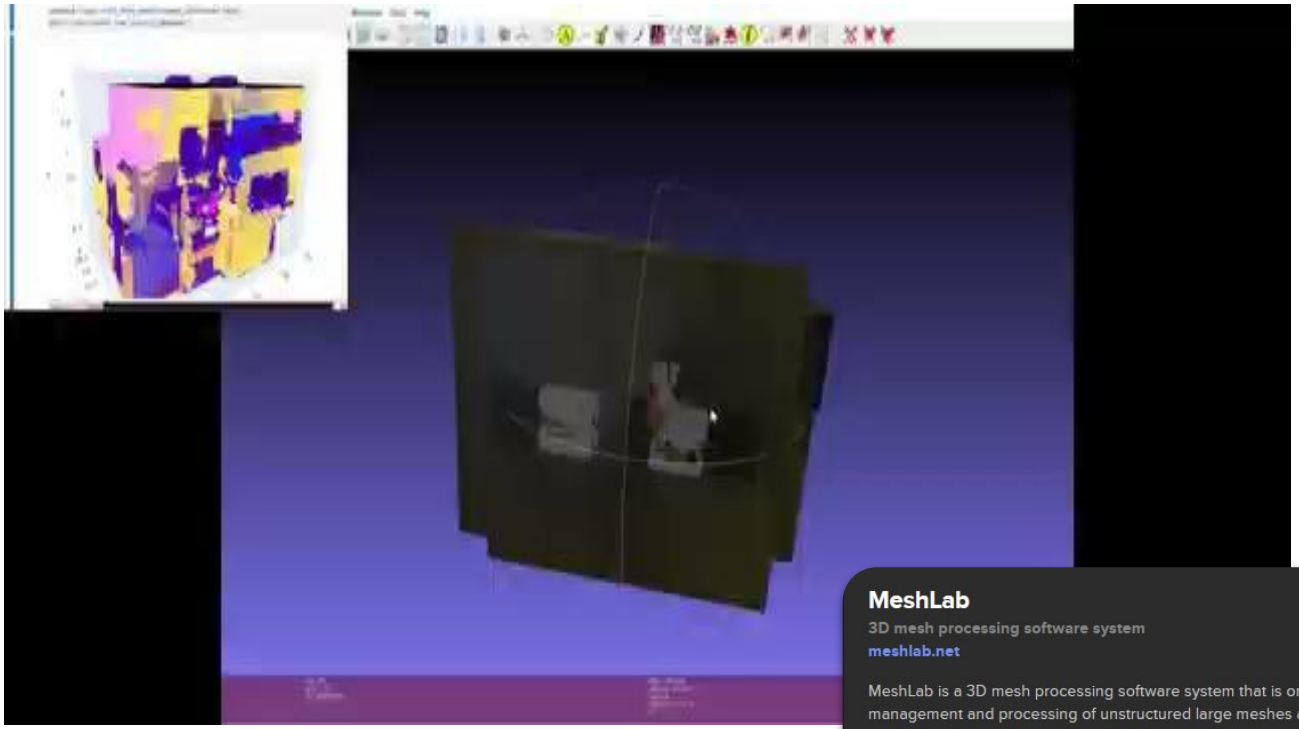
Filter: WindowNorm | Transformer | Multi-modal | multi-view | GCN | untagged

Edit Leaderboard

		Mean IoU	mAcc	oAcc	FLOPs	Number of params	mIoU	Params (M)	Data			
1	PTv3 + PPT	80.8	87.7	92.6		24.1M			✓	Point Transformer V3: Simpler, Faster, Stronger		2023
2	PonderV2 + SparseUNet	79.9	86.5	92.5					✓	PonderV2: Pave the Way for 3D Foundation Model with A Universal Pre-training Paradigm		2023
3	Swin3D-L	79.8	88.0	92.4		N/A			✓	Swin3D: A Pretrained Transformer Backbone for 3D Indoor Scene Understanding		2023

Compiling C for custom cuda components…is not for the faint hearted

Models are…not invariant under transformations



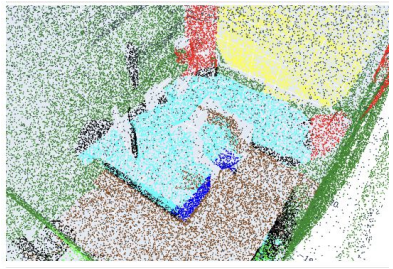
MeshLab

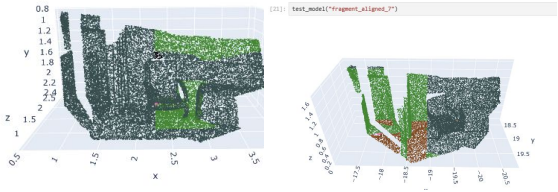
3D mesh processing software system

meshlab.net

MeshLab is a 3D mesh processing software system that is oriented to the management and processing of unstructured large meshes and provides a set of tools for editing, cleaning, healing, inspecting, rendering, and converting these kinds of meshes. Wikipedia

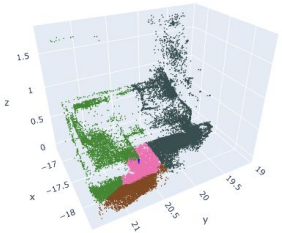
PointNet



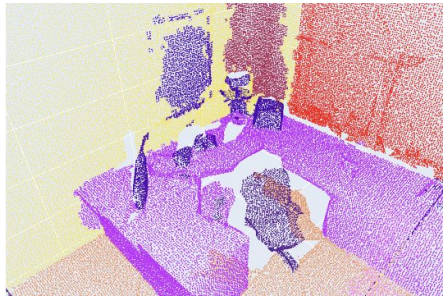


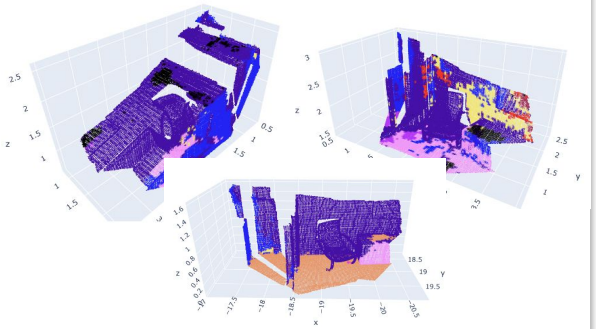
```
[21]: test_model("fragment_aligned_7")
```

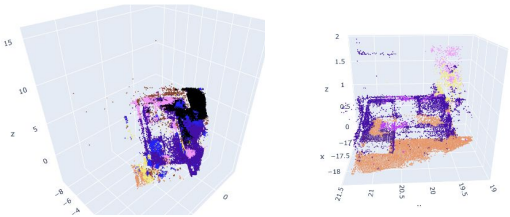
```
: test_model("sofa_aligned")
```



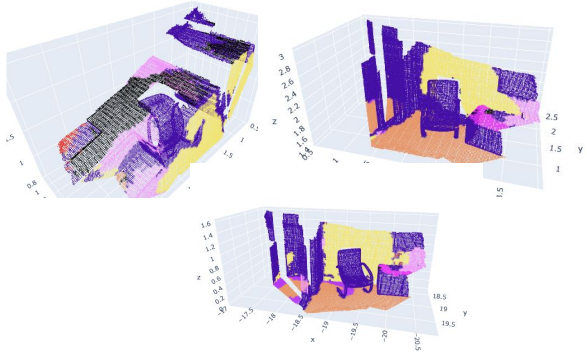
RandLANet

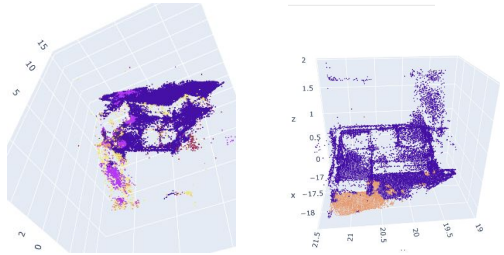




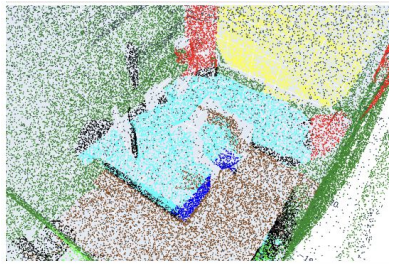


PointTransformer

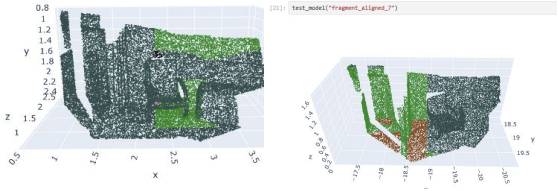


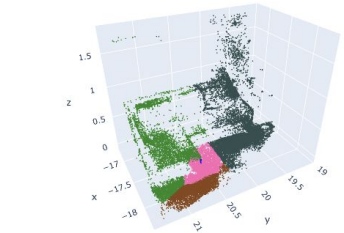


PointNet

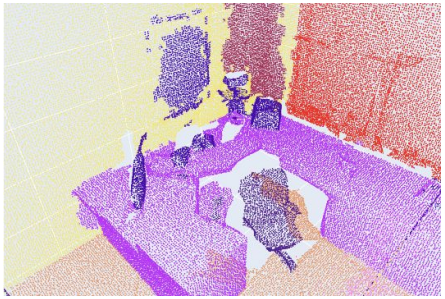


Aligned, scaled, transposed

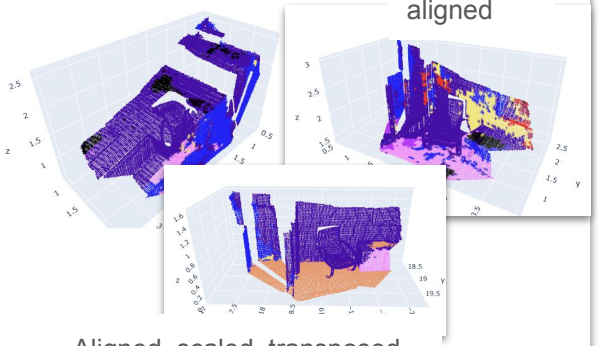




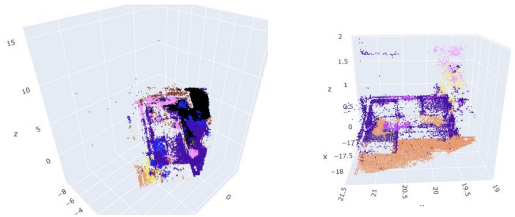
RandLANet



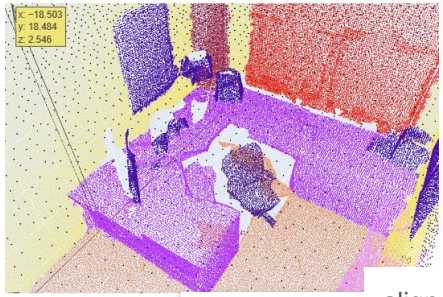
aligned



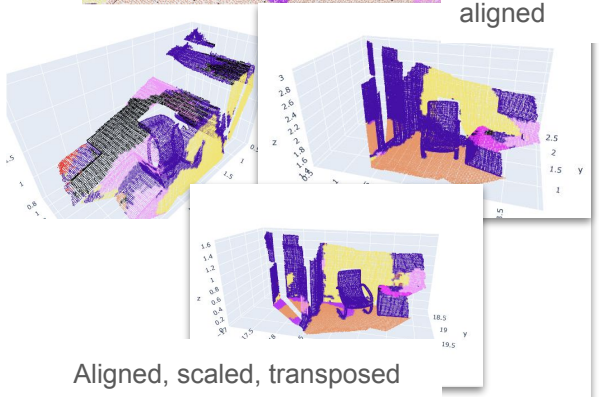
Aligned, scaled, transposed



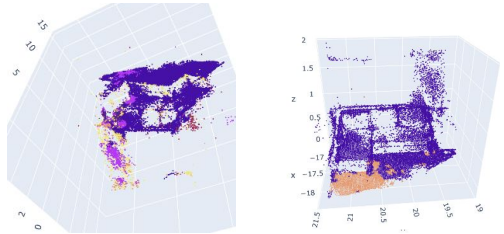
PointTransformer



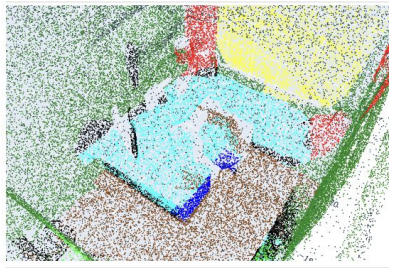
aligned

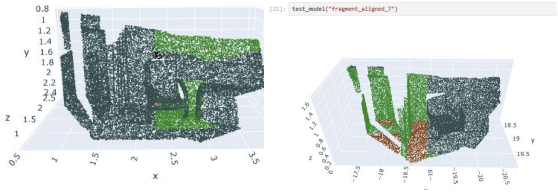


Aligned, scaled, transposed

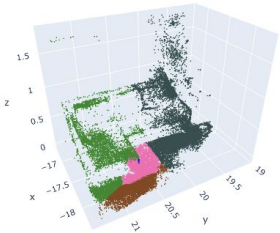


PointNet

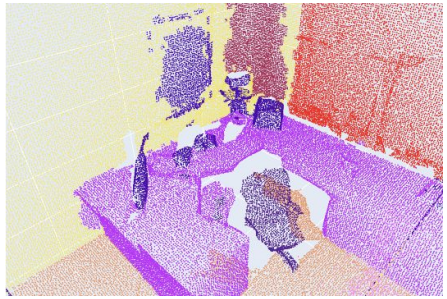


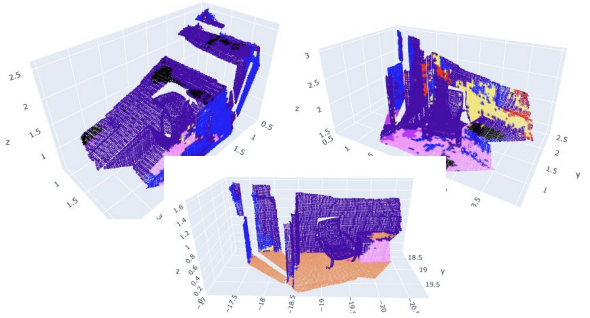


```
test_model("sofa_aligned")
```

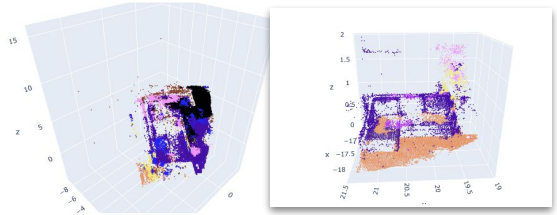


RandLANet

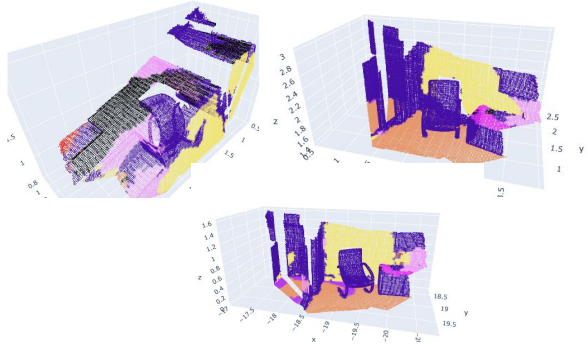




Aligned, scaled, transposed



PointTransformer



Aligned, scaled, transposed