Extracting Levels of Detail of Indoor Spaces using Deep Learning for Indoor Cartography

Enbo Zhou, Keith C. Clarke, Haoyu Shi, Joel Salzman, Bryn Morgan
Department of Geography

University of California, Santa Barbara

Email: ebzhou@geog.ucsb.edu

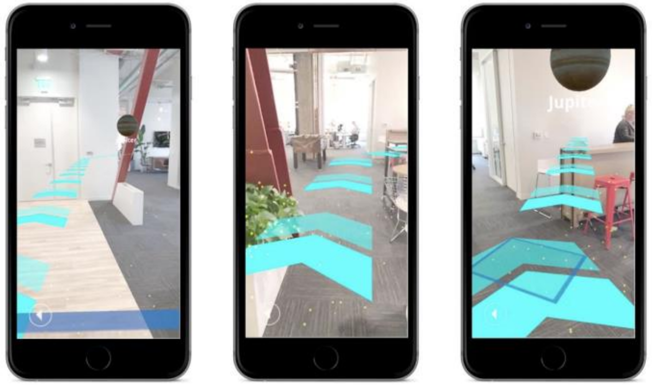
Nov. 18, 2020
AutoCarto @ Redlands, CA, USA

Outline

- Introduction
- Methods
- Experiments, Analysis, and Results
- Conclusions

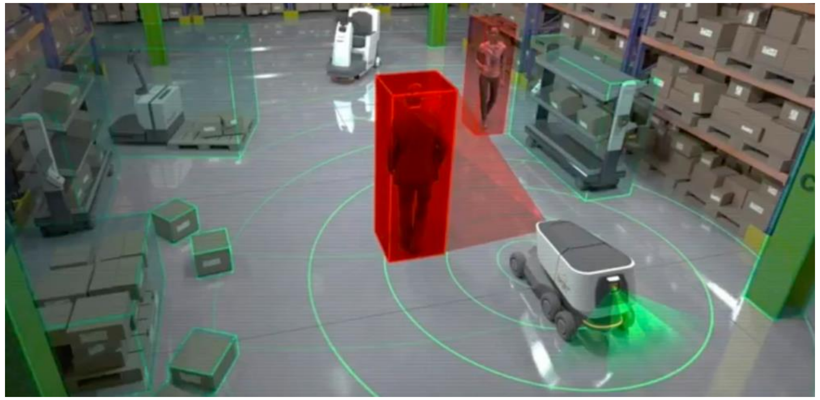
Background

Many applications call for the rapid production and update of indoor maps.



Indoor navigation (mapbox)

Building information management (lodplanner)



Indoor robot

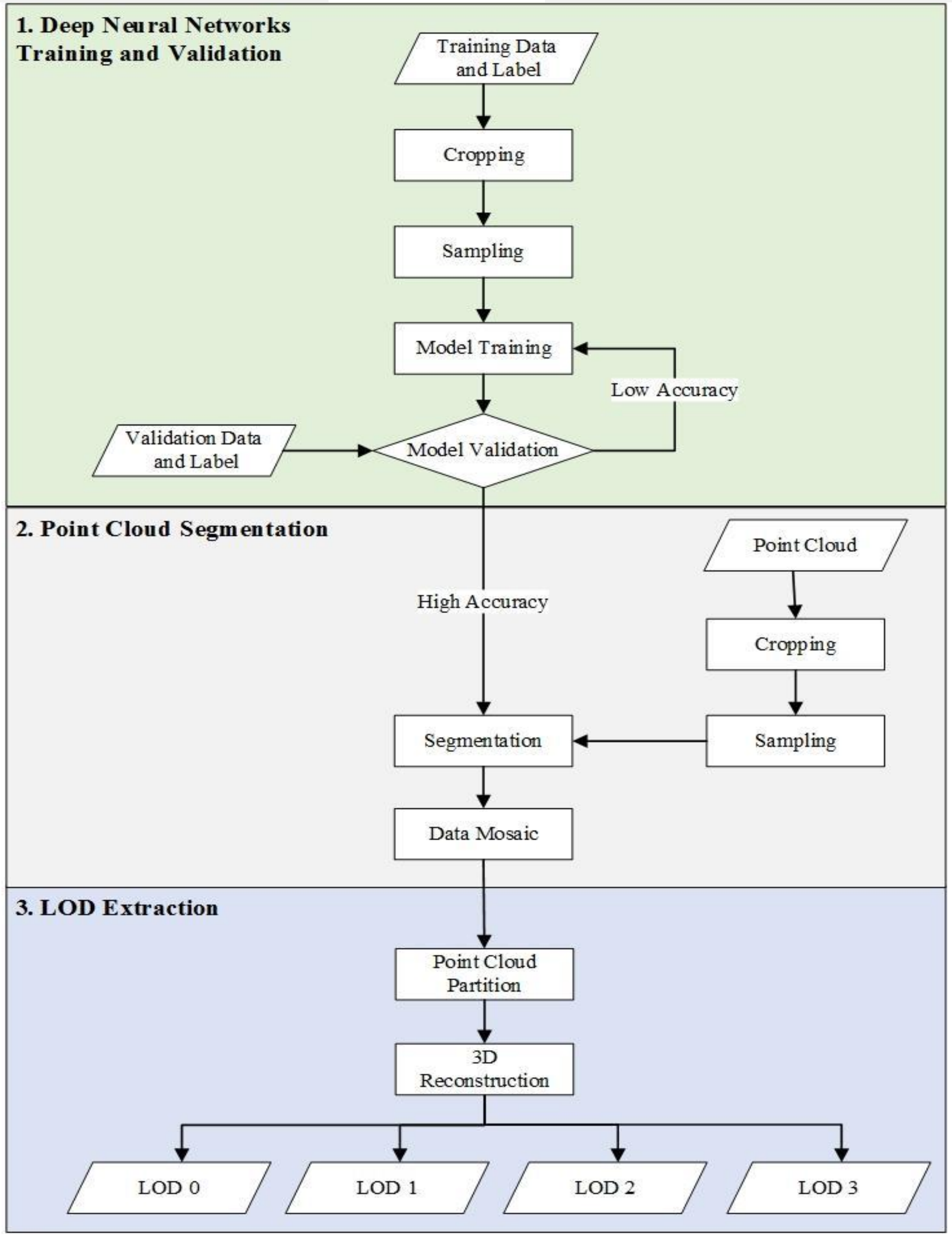
https://www.therobotreport.com/autonomous-navigation-design-challenges/

Introduction

- LiDAR is a popular means for indoor mapping data collecting.
- A huge conversion gap between raw LiDAR point cloud and 3D model exists.
- Traditional methods are inaccurate, and often involve human control.
- No consistent standards for indoor mapping.
- In this research, we designed a deep learning-based framework to extract levels of detail of indoor spaces from raw LiDAR data.

The Proposed Framework

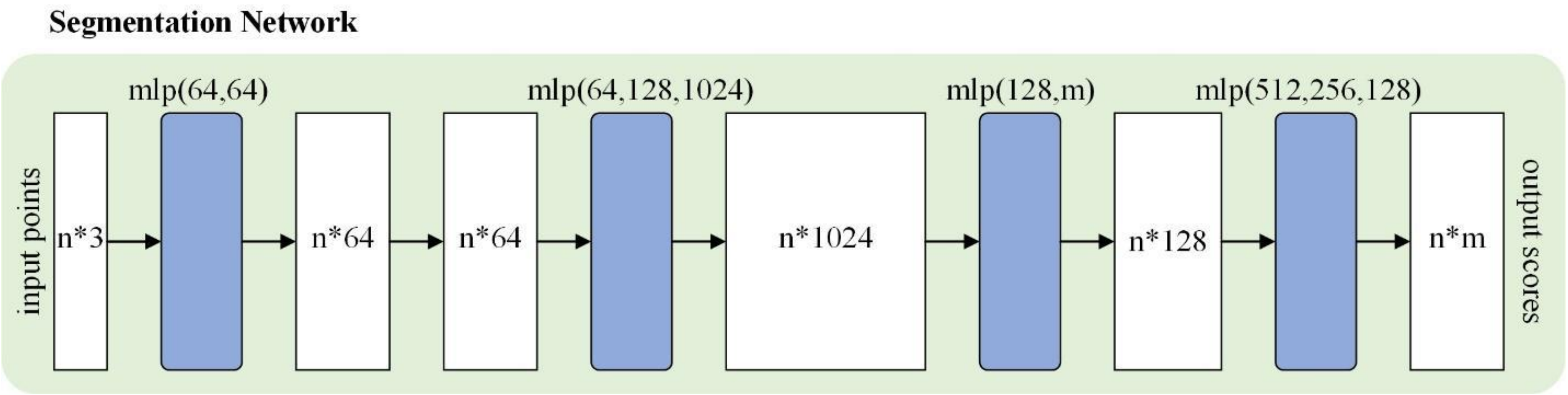
- The proposed framework includes:
  - ➢ Neural network model training
  - ➢ Point cloud segmentation
  - ➢ LOD Extraction
- Step 1: Train a neural network for point cloud segmentation using labelled data.
- Step 2: Segment an experimental point cloud.
- Step 3: Reconstruct 3D LODs from segmented point cloud.



Deep Learning Models

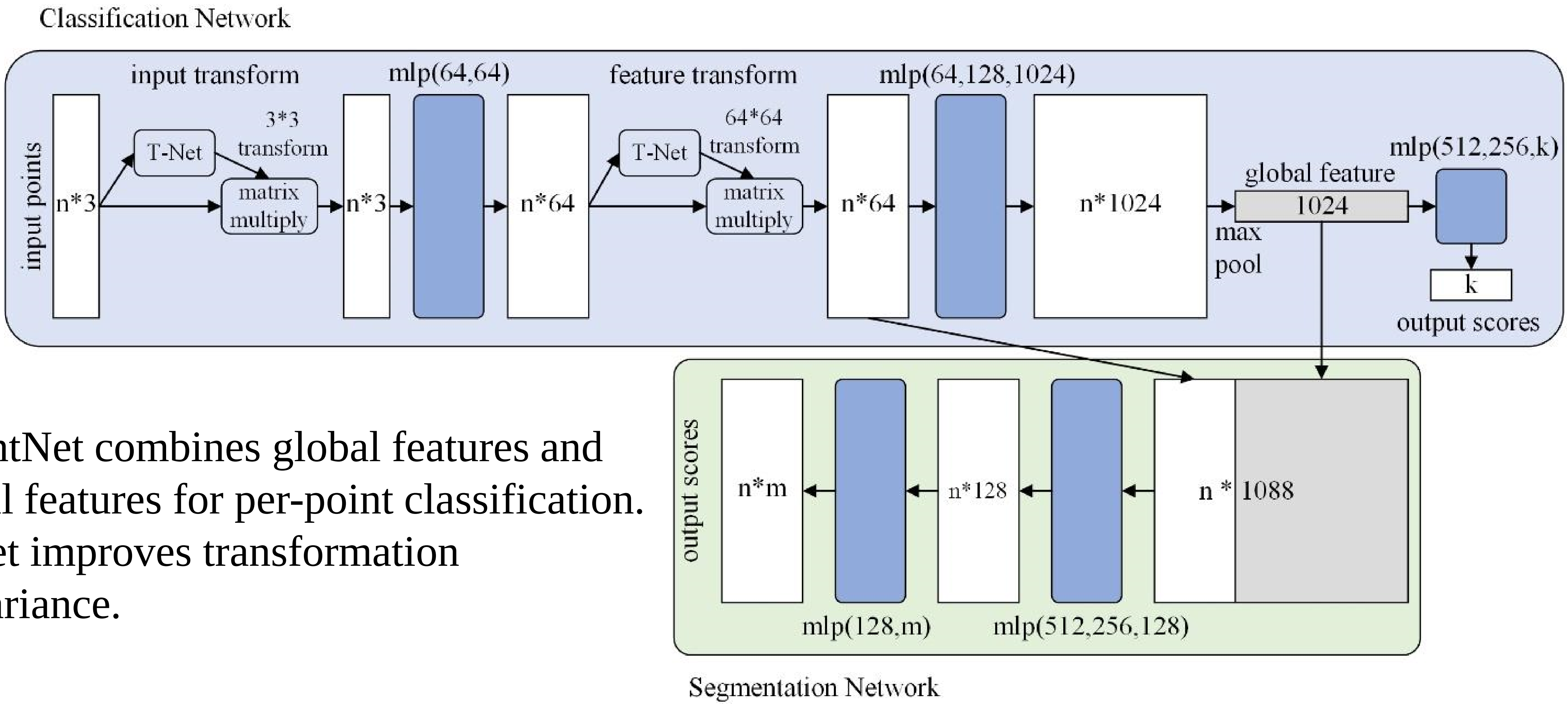
- Neural network models are used in this framework for point cloud segmentation.
- Three neural networks were tested in this research:
  - ➢Vanilla PointNet
  - ➢PointNet (Qi et al., 2017a)
  - ➢PointNet++ (Qi et al., 2017b)
- LODs are reconstructed from a segmented point cloud.
- Used Poisson reconstruction (Kazhdan et al., 2006; Kazhdan and Hoppe, 2013).

Vanilla PointNet



- A simplified PointNet network consists of multilayer perceptrons for point cloud segmentation.
- Used as a baseline for the comparison of different segmentation methods.

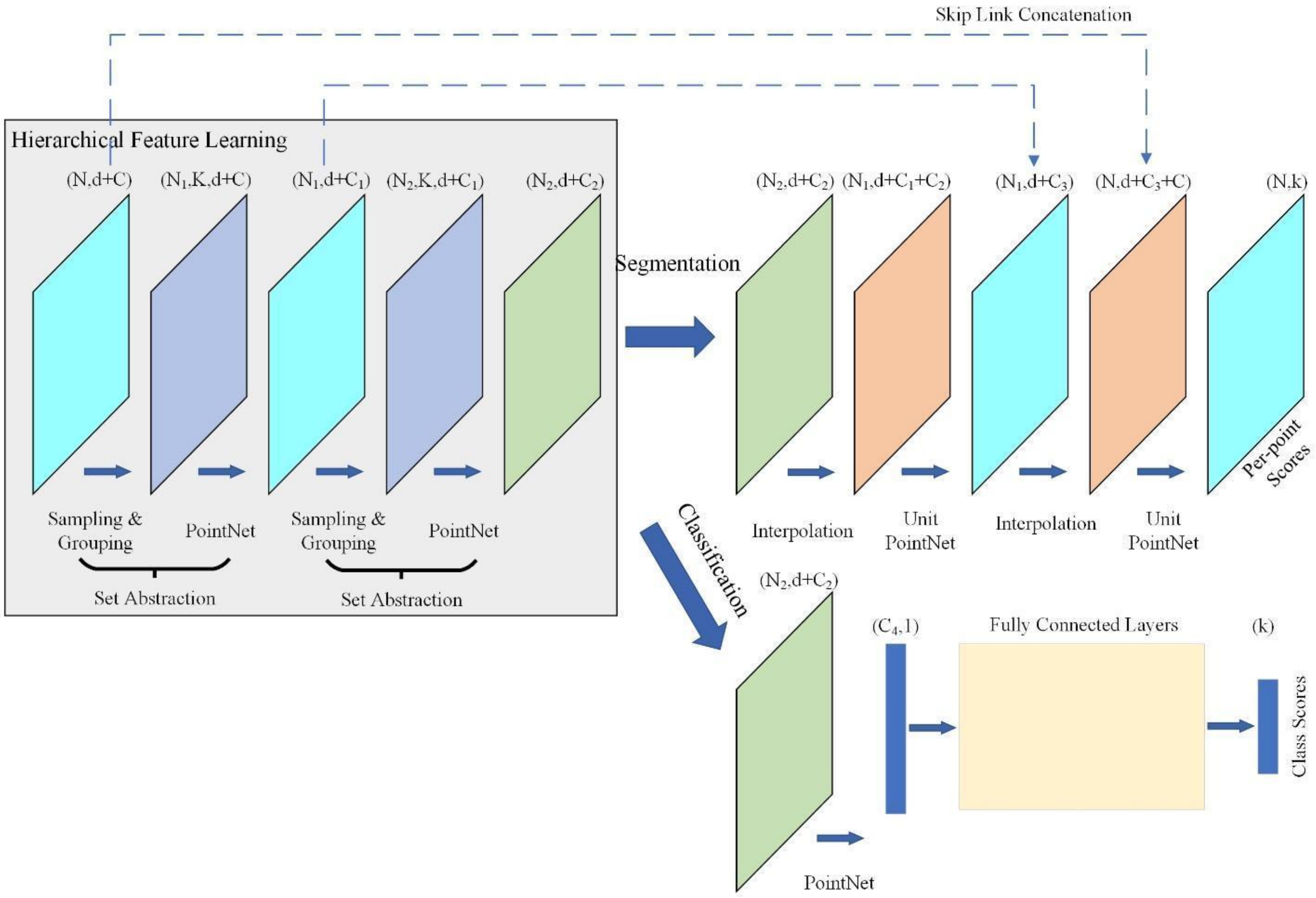
PointNet



PointNet combines global features and local features for per-point classification. T-net improves transformation invariance.

Qi et al., 2017a

PointNet++

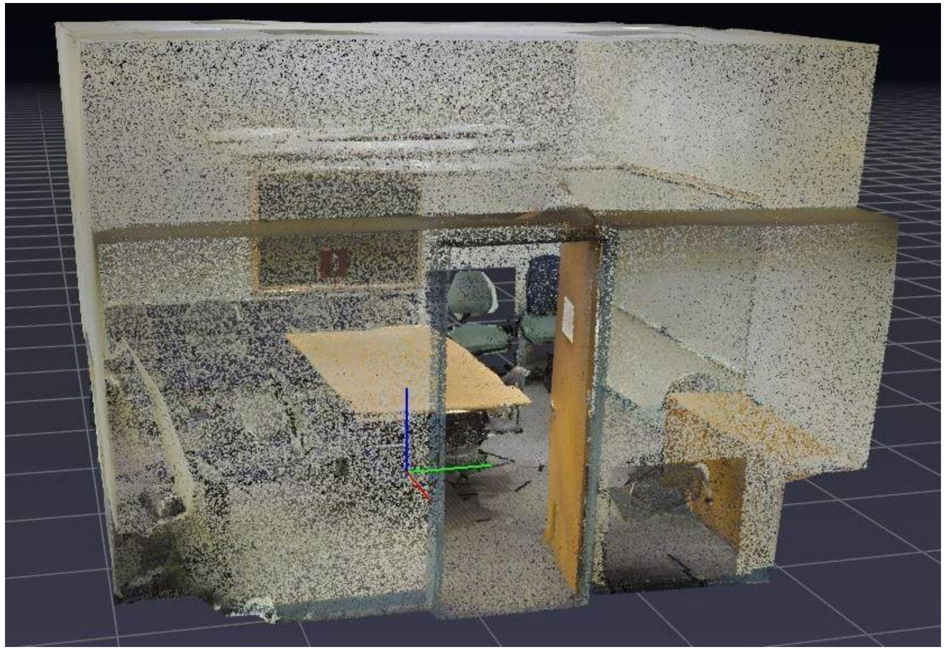


- PointNet++ follows the encoder-decoder structure, and takes the effects of neighboring points and varying point density into account, which makes the method more robust and scalable.
- Designed for both point cloud classification and segmentation.

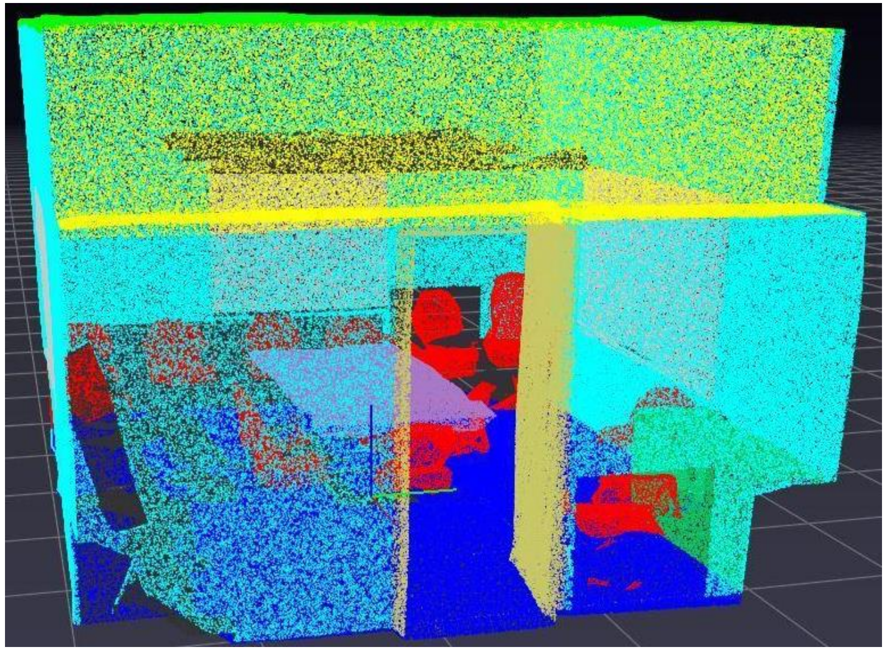
Qi et al., 2017b

Experimental Data

- Stanford Large-scale 3D Indoor Spaces Dataset (S3DIS, Armeni et al. 2016) for model training.
  - 13 semantic categories including structural elements (ceiling, floor, wall, beam, column, window and door), common indoor objects (table, chair, sofa, bookcase and board) and clutter.



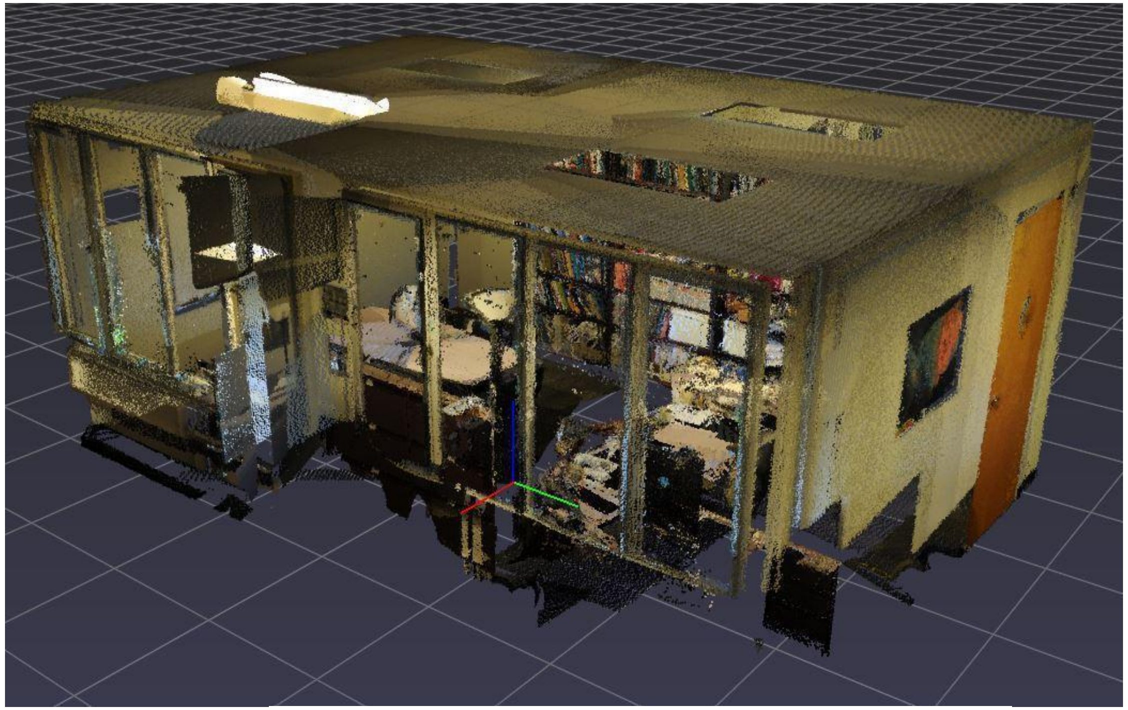
Raw point cloud of a meeting room



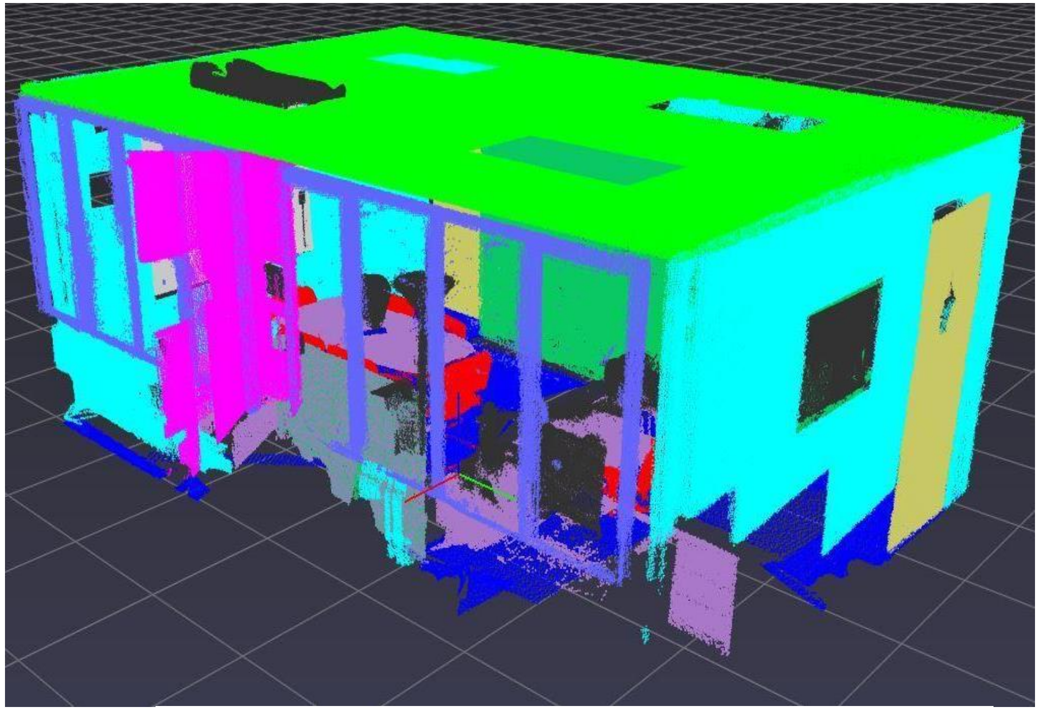
Labelled point cloud of a meeting room

UCSB Office Data

- LiDAR point cloud for UCSB Ellison Hall Room 1720.
- Five scans were merged into a point cloud, which includes 5,306,422 data points. The same 13 categories as SIDIS.

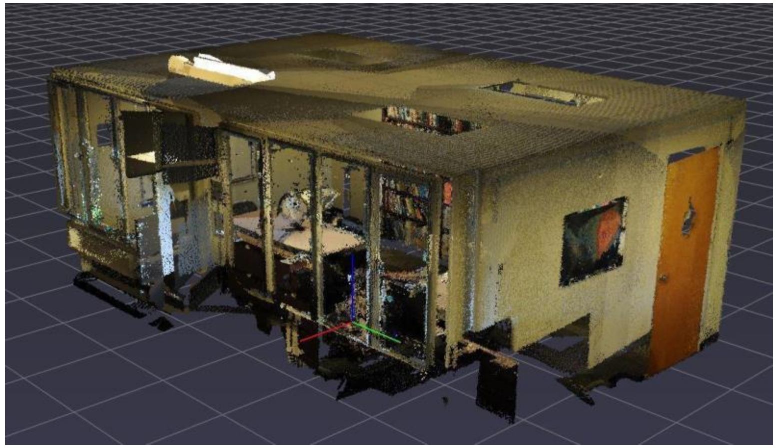


Raw point cloud

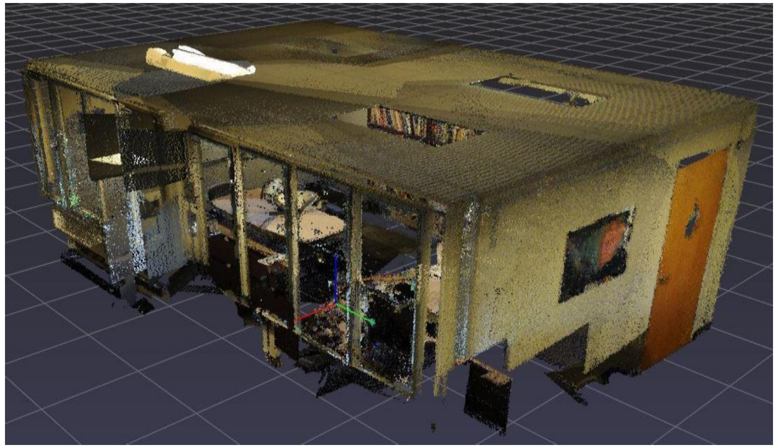


Labelled point cloud

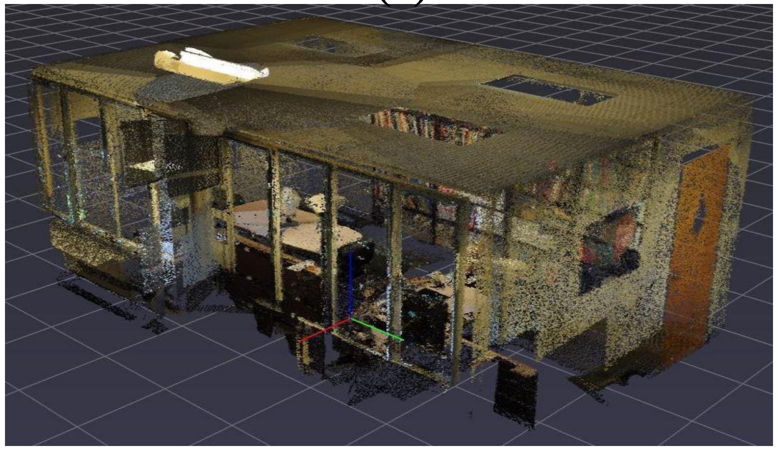
Data Cropping and Subsampling



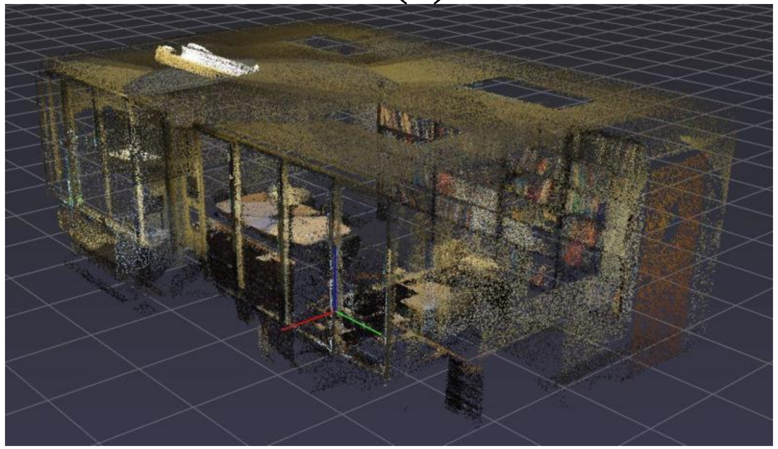
(a)



(b)



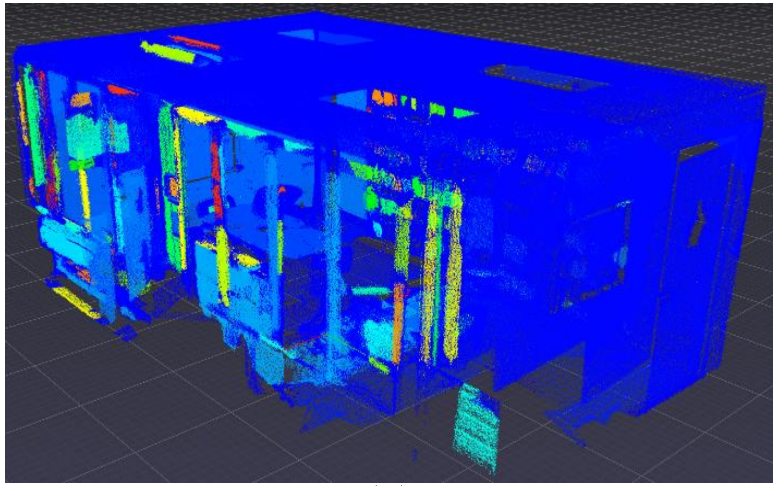
(c)



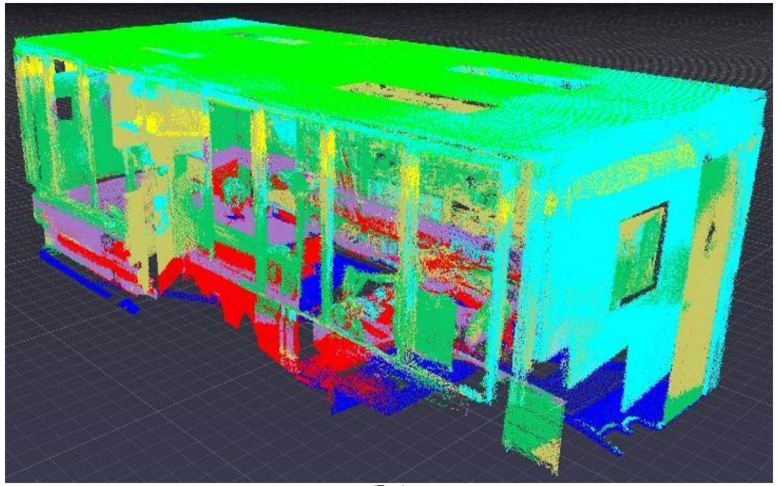
(d)

The Effect of the Sampling Block Size and Stride on Sampled Point Cloud Density: (a): 0.1-m sampling block and 0.05-m stride, (b): 0.25-m sampling block and 0.125-m stride, (c): 0.5-m sampling block and 0.25-m stride, (d): 1-m sampling block and 0.5-m stride.

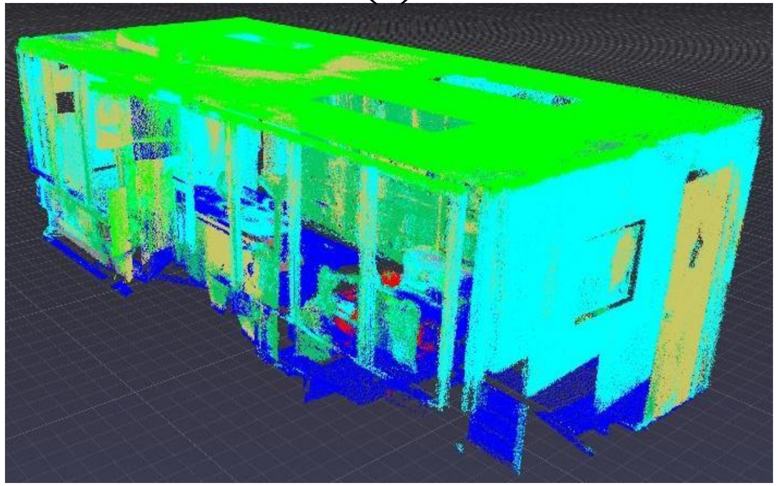
Sample (a) Segmentation Results



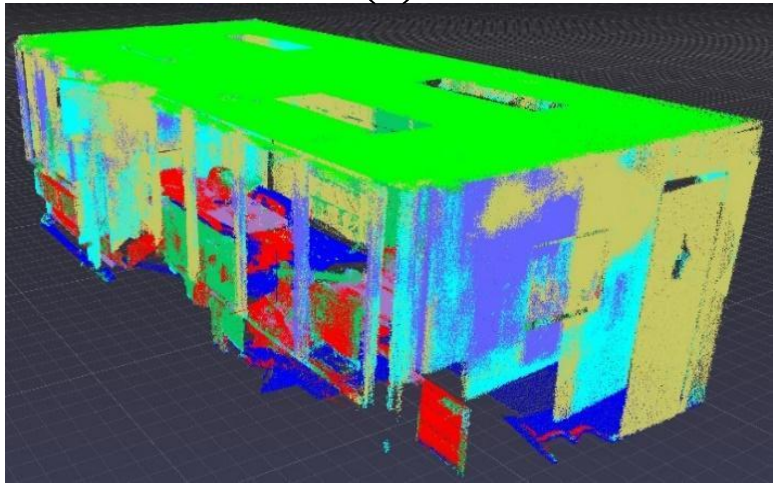
(a)



(b)



(c)



(d)

Segmentation results of region growing, Vanilla PointNet, PointNet and PointNet++ for **sampled data (a) with 0.1-m sampling block and 0.05-m stride**.

Segmentation Results Evaluation

Category	**Sample (a) Accuracy**			Sample (b) Accuracy			Sample (c) Accuracy			Sample (d) Accuracy		
	VPT	PT	**PT++**	VPT	PT	PT++	VPT	PT	PT++	VPT	PT	PT++
Ceiling	80.85%	90.96%	**98.82%**	81.82%	98.98%	99.05%	82.73%	97.08%	99.50%	83.42%	95.99%	99.75%
Floor	92.56%	99.73%	**98.07%**	87.12%	99.66%	99.53%	88.88%	99.13%	99.31%	89.52%	98.60%	98.95%
Wall	42.01%	69.50%	**30.75%**	51.73%	69.71%	24.86%	45.63%	50.02%	19.70%	72.32%	39.85%	28.04%
Column	0.27%	0.00%	**0.00%**	0.39%	0.00%	0.00%	0.46%	0.00%	0.00%	0.41%	0.00%	0.00%
Window	0.03%	0.03%	**15.67%**	0.08%	0.07%	15.47%	0.07%	0.50%	5.30%	0.08%	0.06%	1.79%
Door	44.45%	55.92%	**94.52%**	53.47%	64.53%	96.82%	43.24%	49.23%	91.57%	41.05%	44.42%	80.76%
Table	69.62%	0.75%	**42.24%**	62.46%	11.33%	54.86%	58.50%	45.50%	67.73%	58.78%	62.49%	65.07%
Chair	50.78%	9.62%	**94.84%**	47.67%	5.59%	75.42%	46.72%	26.76%	65.47%	45.66%	74.37%	67.40%
Bookcase	45.27%	41.01%	**64.18%**	42.75%	65.92%	79.12%	38.84%	81.82%	81.74%	35.67%	57.11%	79.13%
Board	0.00%	0.00%	**0.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%
Clutter	0.00%	0.00%	**0.00%**	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total	67.84%	72.47%	**80.43%**	64.13%	74.53%	76.53%	57.40%	68.80%	70.65%	51.43%	56.64%	60.38%
Kappa	0.5676	0.5248	**0.6474**	0.5329	0.6266	0.6470	0.4895	0.5998	0.5941	0.4152	0.4750	0.5074

VPT: Vanilla PointNet
PT: PointNet
PT++: PointNet++

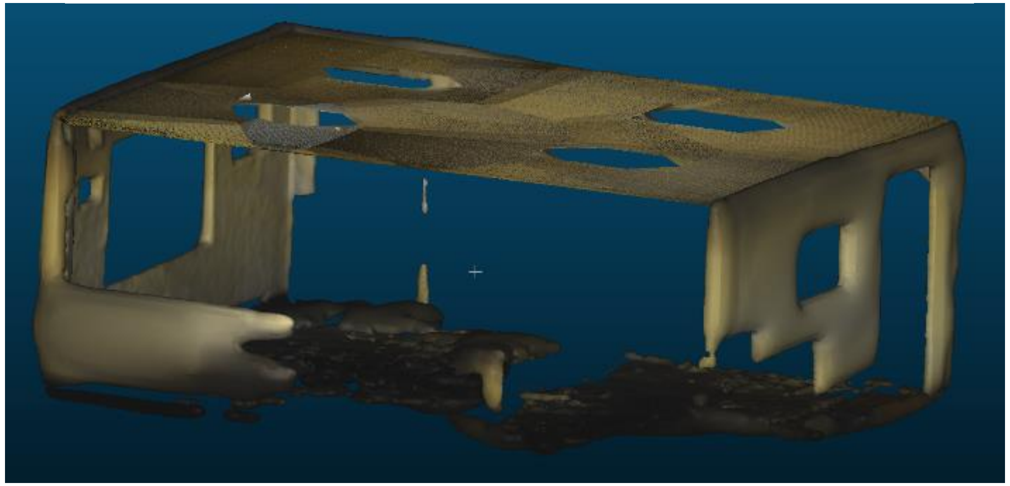
Indoor LOD Standard Design: Based on Object Permanence

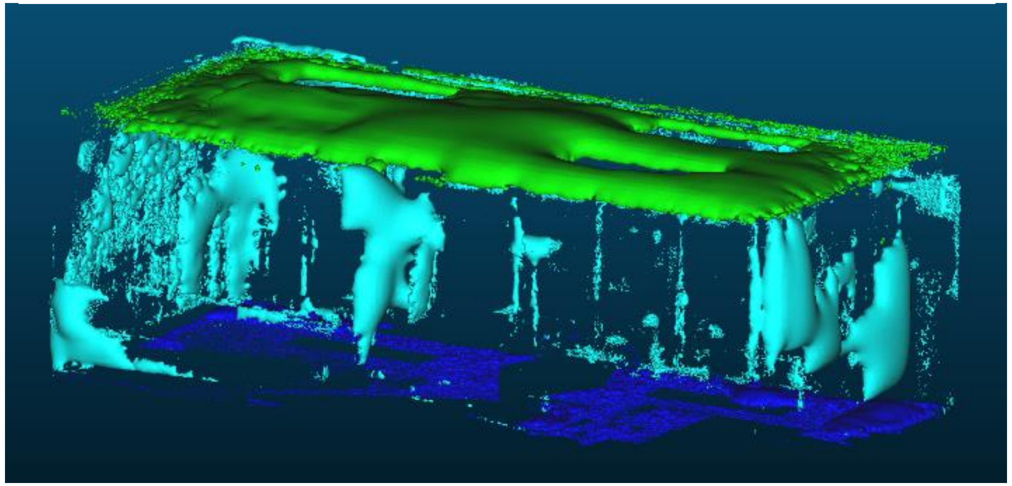
LOD	Objects
0	Wall, ceiling, floor. (Most permanent)
1	LOD0 + column, door, window. (Permanent)
2	LOD1 + table, case, board. (Semi-permanent)
3	LOD2 + chair + clutter (Transient)

LOD Reconstruction Results

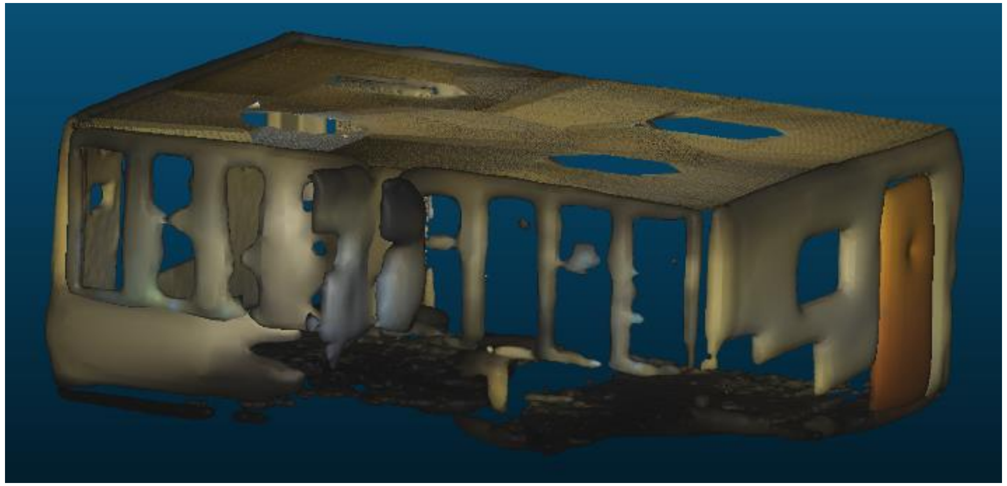
Manual segmentation + Poisson Reconstruction | **PointNet++ segmentation + Poisson Reconstruction**

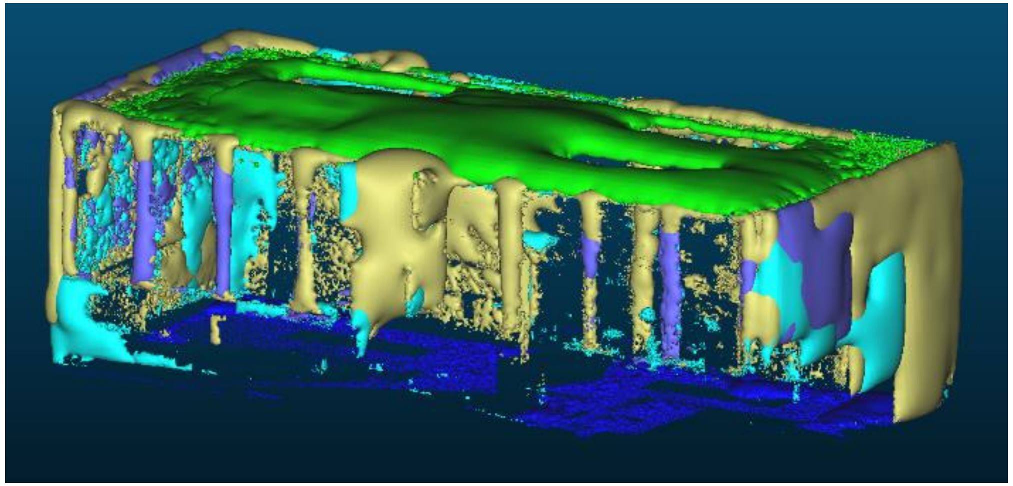
LOD 0



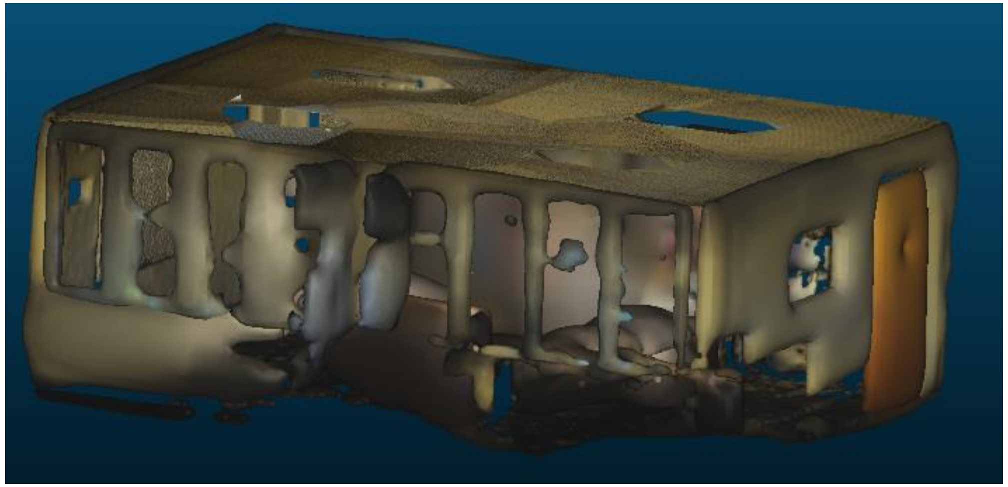
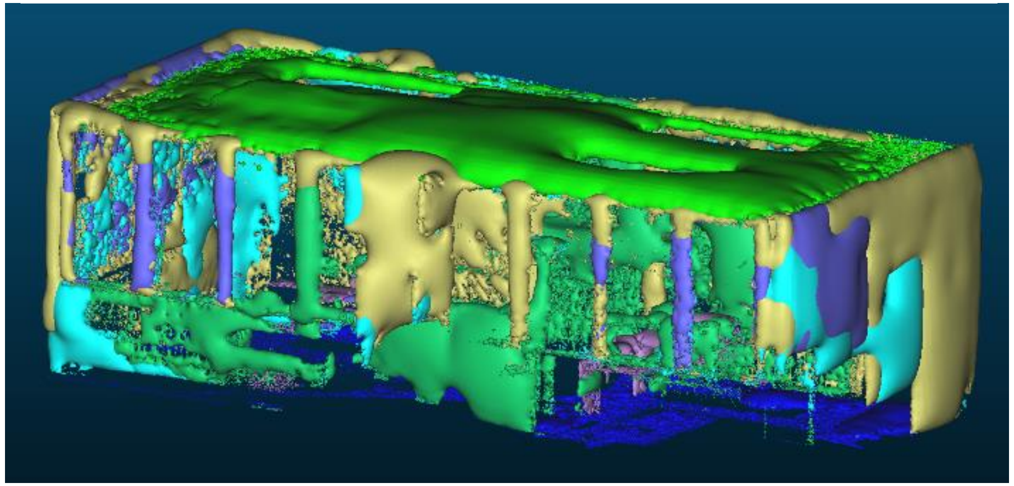
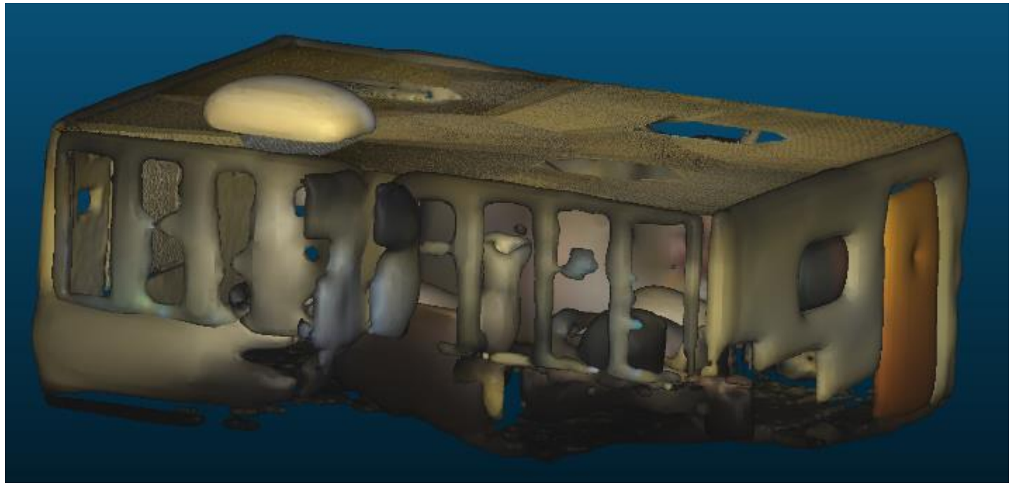
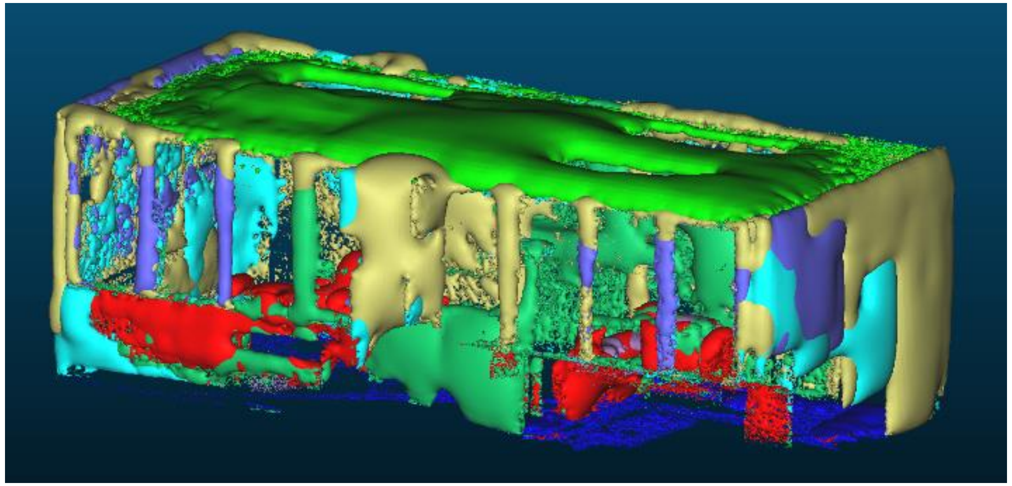


LOD 1





LOD Reconstruction Results

	Manual segmentation + Poisson Reconstruction	PointNet++ segmentation + Poisson Reconstruction
LOD 2		
LOD 3		

LOD Accuracy

Model	LOD0		LOD1		LOD2		LOD3	
	Mean	**Standard Deviation (m)**	**Mean (m)**	**Standard Deviation (m)**	**Mean (m)**	**Standard Deviation (m)**	**Mean (m)**	**Standard Deviation (m)**
Manual segmentation + Poisson Reconstruction	0.0284	0.0589	0.0266	0.0493	0.0320	0.0469	0.0327	0.0450
PointNet++ segmentation + Poisson Reconstruction	0.0315	0.0489	0.0434	0.0581	0.0476	0.0568	0.0482	0.0544

LODs produced from neural networks achieve competitive accuracy to those from manual modelling.

Discussion

- Point cloud segmentation is critical in the framework. Different segmentation methods influence following LOD reconstruction accuracy.
- The three deep neural network models are much better than region growing. PointNet++ performed best in our experiments.
- Point density also affects segmentation. Higher point density contributes to overall segmentation accuracy.
- Neural networks failed to identify some objects, especially the board, columns and windows. More information such as calibrated intensity, reflectance, etc. may be needed.
- A multiple-step segmentation may help differentiate similar objects.

Conclusions

- We designed a framework to generate LODs from raw point cloud.
- Explored applying state-of-art deep neural network models to segment LiDAR point clouds.
- Segmentation results were aggregated into four levels of detail according to our ad hoc indoor cartography standard.
- Experimental results demonstrate the feasibility of the proposed framework.
- Future studies will consider more efficient segmentation methods, 3D object reconstruction from local incomplete point clouds, and experiments on a larger building interior
- A long-term goal is a scan-to-model computational workflow, that serves as a primary mapping tool for indoor cartography.

Thank you for your attention!

Q&A

ebzhou@geog.ucsb.edu