הפקולטה להנדסת חשמל ומחשבים ע"ש אנדרו וארנה ויטרבי

SIPL Signal and Image Processing Lab.

Technion Israel Institute of Technology

3D Object Detection For Intel RealSense LiDAR

Anaelle Yasdi, Judit Ben Ami, Ori Bryt

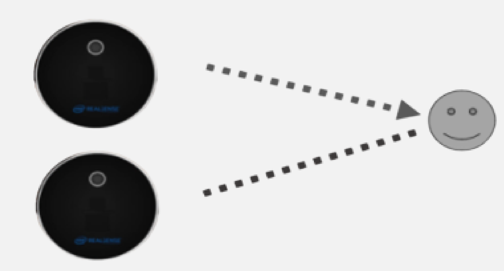
Work In Progress

Introduction

RealSense LiDAR camera – Light Detection And Ranging

Time-Of-Flight depth camera that sends a laser pulse and calculate the z using:

$$z = \frac{c \cdot t}{2} \qquad c - \text{speed of light} \qquad t - \text{time of flight}$$



3D Object Detection:

Estimate the oriented 3D bounding boxes and classes from 3d data

Objective

Object detection and classification from point clouds taken by L515, using deep learning methods

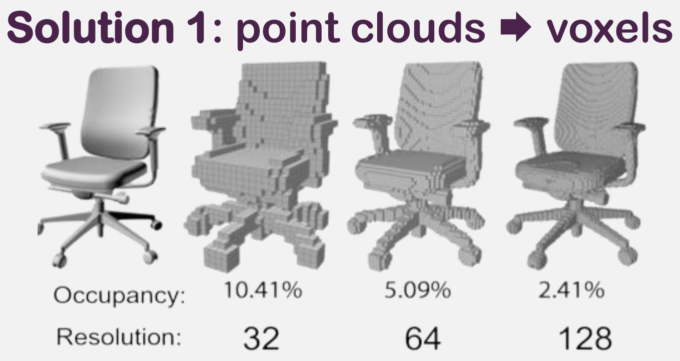
- Real scene images
- Use pre-trained network (VoteNet) and adapting to L515 data using transfer learning

NN: PointNet

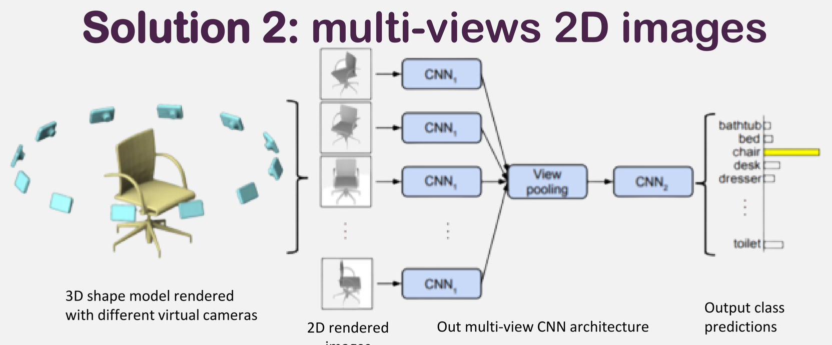
Problem: point cloud is N orderless points, each represented by D dimensional vector

The model needs to be invariant to N! permutation

Solution 1: point clouds ➡ voxels



Solution 2: multi-views 2D images



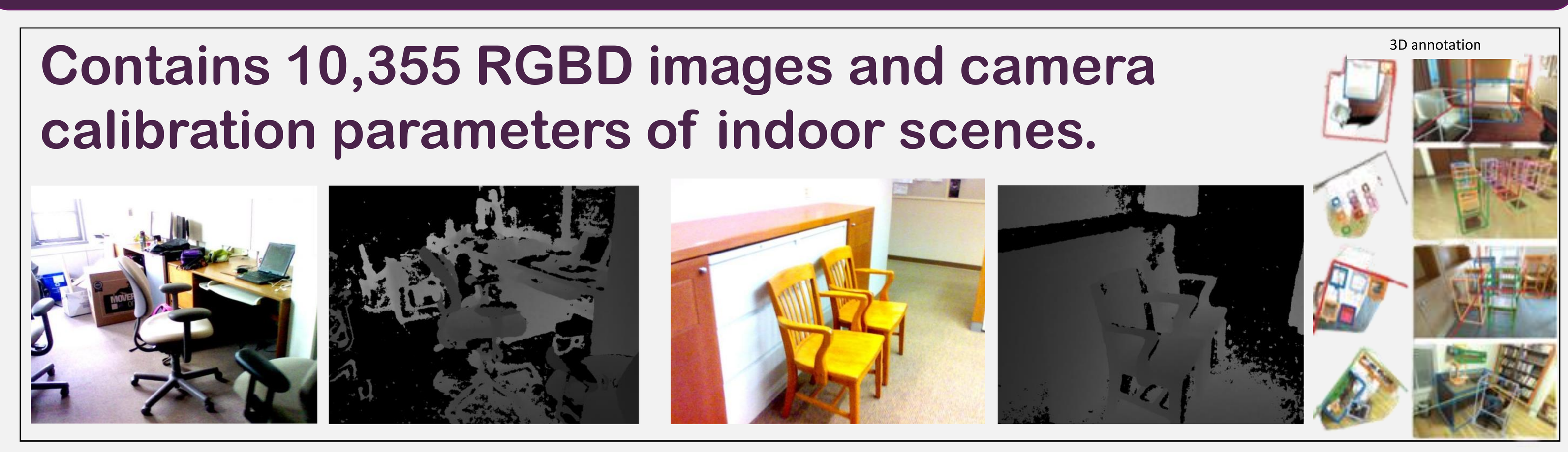
PointNet solution: Use Permutation-Invariant Symmetric function:

$$f(x_1, x_2, \dots, x_n) = \max\{x_1, x_2, \dots, x_n\}$$

$$f(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_n$$

Database: SUNRGBD

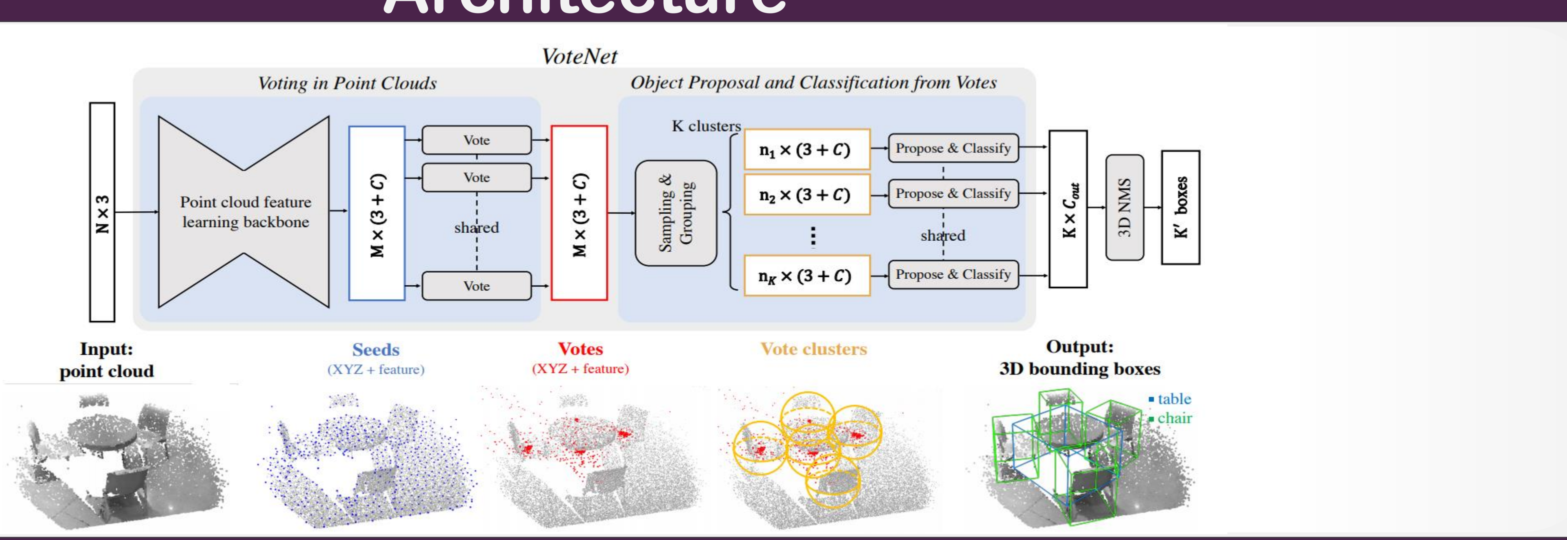
Contains 10,355 RGBD images and camera calibration parameters of indoor scenes.



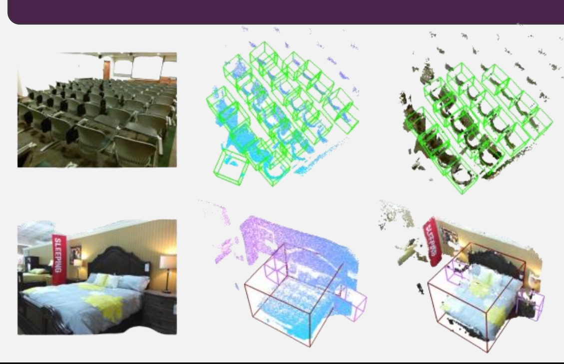
Chosen NN: VoteNet

- The Point Cloud feature learning backbone is based on PointNet++
- From the features extracted, a Voting procedure is done based on Deep Hough voting
- Votes are divided into K clusters by spatial clustering
- PointNet-like network generates the object proposals from votes

Architecture



Results on SUNRGBD



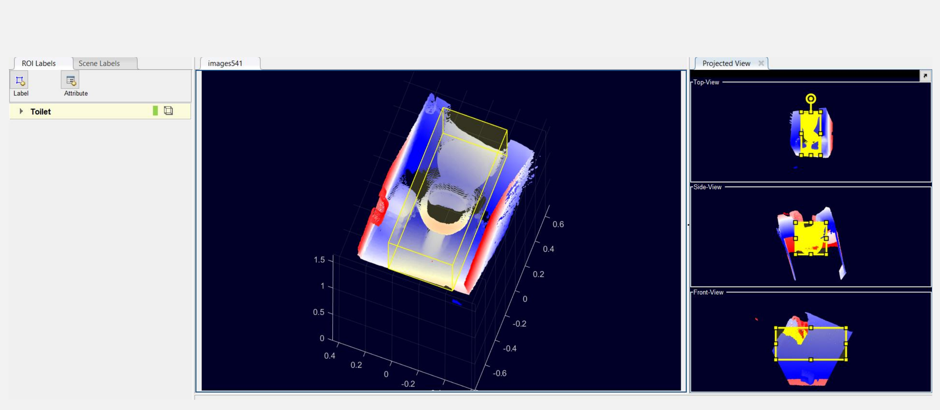
	Input	bathtub	bed	bookshelf	chair	desk	dresser	nightstand	sofa	table	toilet	mAP
DSS [42]	Geo + RGB	44.2	78.8	11.9	61.2	20.5	6.4	15.4	53.5	50.3	78.9	42.1
COG [38]	Geo + RGB	58.3	63.7	31.8	62.2	**45.2**	15.5	27.4	51.0	**51.3**	70.1	47.6
2D-driven [20]	Geo + RGB	43.5	64.5	31.4	48.3	27.9	25.9	41.9	50.4	37.0	80.4	45.1
F-PointNet [34]	Geo + RGB	43.3	81.1	**33.3**	64.2	24.7	**32.0**	58.1	61.1	51.1	**90.9**	54.0
VoteNet (ours)	**Geo only**	**74.4**	**83.0**	28.8	**75.3**	22.0	29.8	**62.2**	**64.0**	47.3	90.1	**57.7**

Solution

Workflow

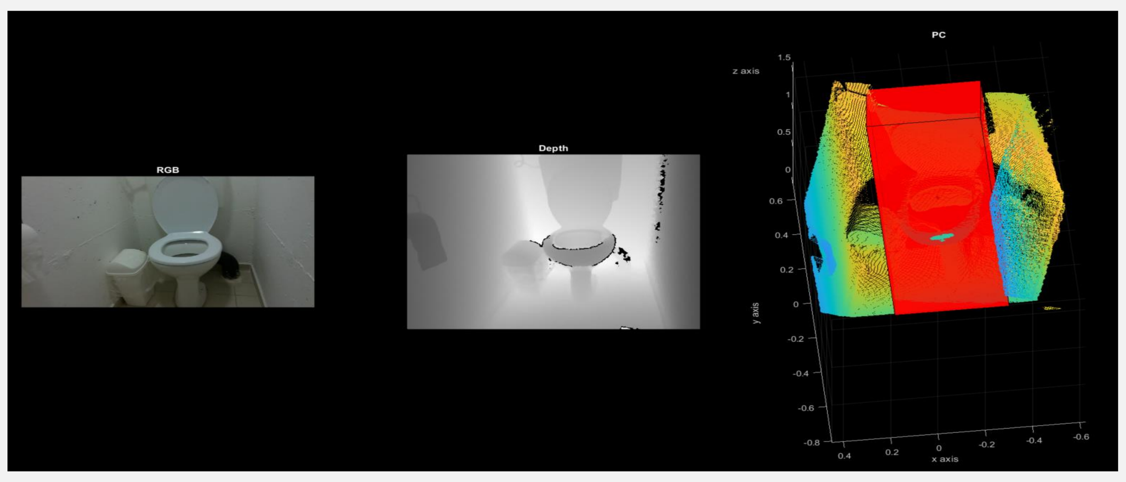
- Take pictures with Real Sense LiDAR L515 camera
- Preprocessing and Image labeling
  - Clean Raw Data (Depth Image) → Point Cloud → Label (Matlab LiDAR ToolBox) → Transform to SUNRGBD-like Data
- Train NN
  - LiDAR L515 (RealSense) Data → Transfer learning
  - SUNRGBD Data → Trained NN → Transfer learning
- Check Test Set
- Accuracy

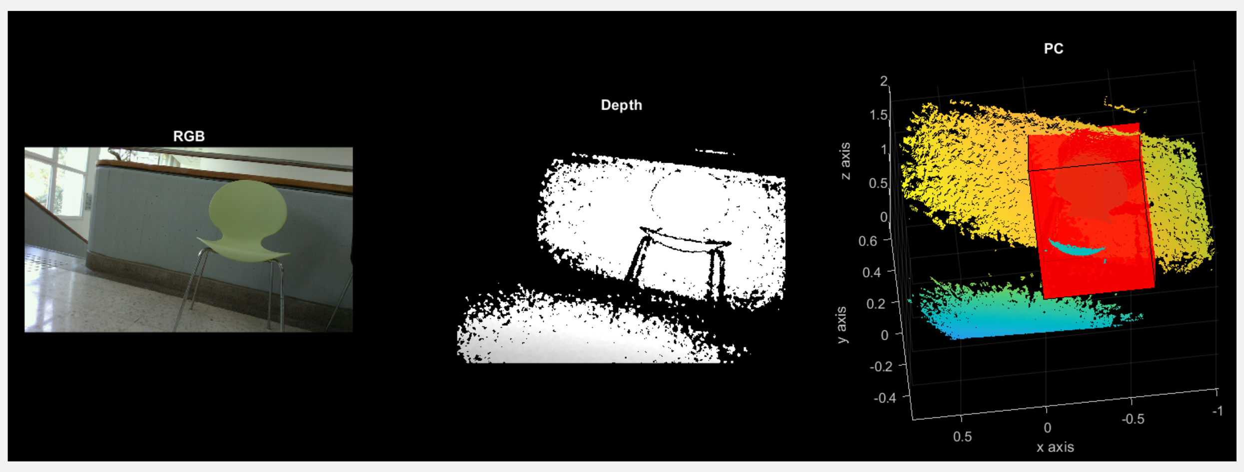
ToolBoox capture



- Matlab ToolBox: Lidar labeler

L515 Database





Conclusions

- The camera doesn't detect well dark and shiny objects
- VoteNet results vary compared to other networks – some classes (chair, toilet) have impressive results, while others have poor results