Plan for 3D Deep

Jungwon Kang, Maryam Jameela, Razieh Ramak

Sept 30 2018

Objectives

- Building 1st Version Deep Network for Each Task

Task	Major Contributor	Objectives
Noise filtering (for Optech)	Razieh Ramak	Point cloud segmentation (Noise/Non-noise), non real-time
Point cloud segmentation (for Optech)	Maryam Jameela	Point cloud segmentation (N-class objects), non real-time
3D object detection (for Thales)	Jungwon Kang	Real-time 3D object detection

Schedule

Month	Task	Deliverable
Oct 2018	• Problem definition • Dataset preparation • Literature survey	• Document describing problem definition, dataset, and literature survey • Visualization of dataset
Nov	• Practicing deep library • Design & implementation	• Document describing design
Dec	• Implementation	• Source code (Dec 31)
Jan 2019	• Documentation	• Document describing implementation (Jan 15)

*Submission deadline of major conferences starts from March.

Management Policy

- Regular meeting or discussion biweekly

- Team website:
  - https://github.com/yorku-ausml/deep3d

Written on Oct 1 2018

To-do List

- Problem definition, including
  - Cause of noise (Razieh)
  - Object classes (Maryam, Jungwon)
- Dataset description, including
  - Existing Optech airborne dataset (Razieh)
  - Dataset size
  - Current repository
  - Visualization
- Etc
  - Finding point cloud label tool (for making ground-truth)
  - Finding visualization tool

Key Literature

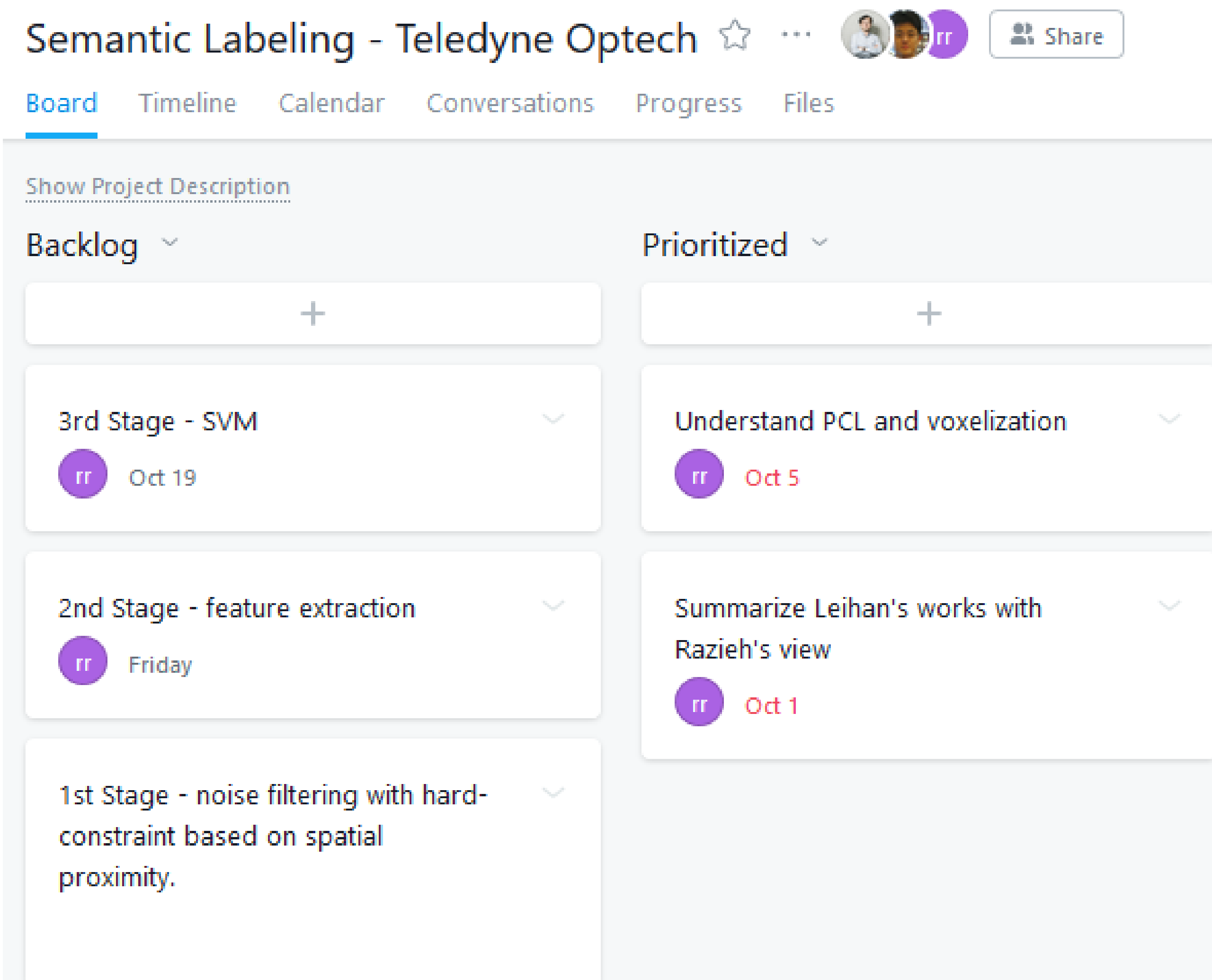
- Point cloud segmentation
  - Large-scale point cloud segmentation with superpoint graphs
    https://github.com/loicland/superpoint_graph
    *Rank 1 in http://www.semantic3d.net/
  - PointNet++: deep hierarchical feature learning on point sets in a metric space
    https://github.com/charlesq34/pointnet2
    *Rank 4 in http://www.semantic3d.net/
- Object detection
  - Joint 3D proposal generation and object detection from view aggregation
    https://github.com/kujason/avod

*Literature list is also available at https://github.com/yorku-ausml/deep3d/wiki/Related-works

Current Progress

Oct 12 2018

Asana Assignment



Progress on Noise Filtering

Razieh

- Atmospheric noise filtering
  - Noise filtering
    Segmentation of raw point cloud using voxelization
    Pre-classification by defining special rules
  - Feature extraction
    Using Eigen library and programming
  - Classification sing SVM
    Using "libSVM"

- Understanding PCL and voxelization
  - PCL
    A large scale, open project for 2D/3D image and point cloud processing. However, there is no PCL in noise filtering application
  - Voxelization
    A data structure used to represent a collection of multi-dimensional points and is commonly used to represent three-dimensional data

Clarifying the Task

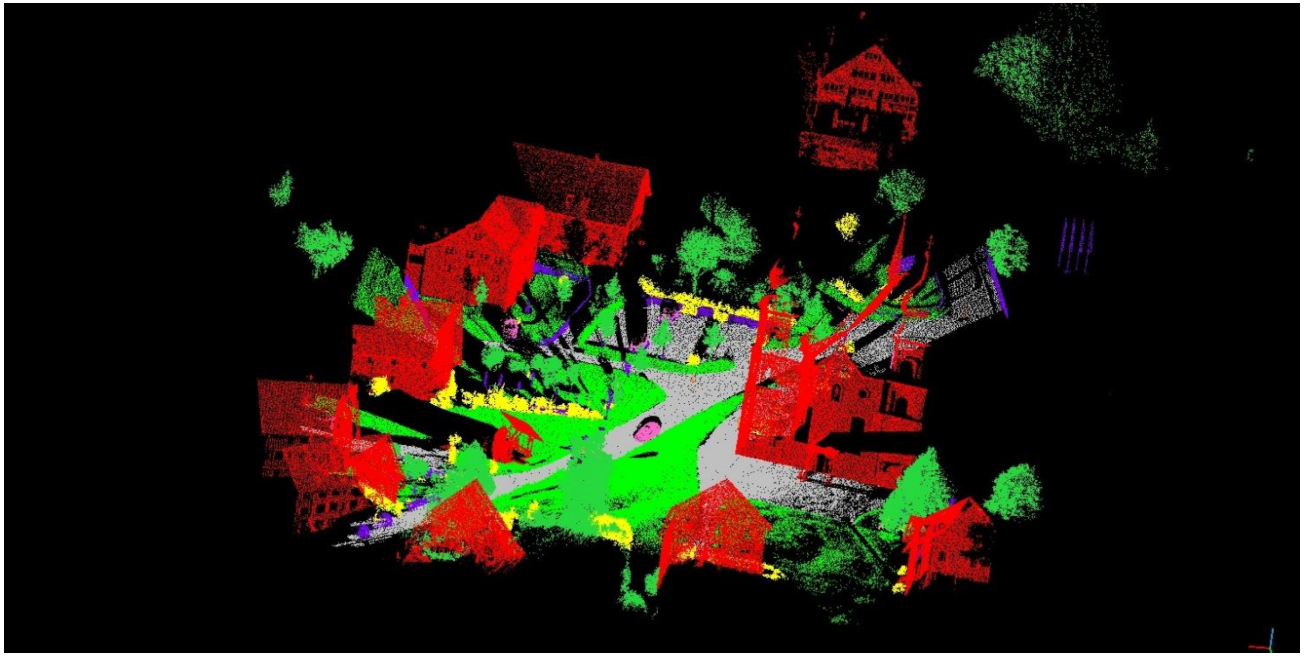
Dataset used for training?
What kind of dataset will be used? Mobile data / airborne lasers / hybrid dataset mixture of both.
What kind of environment? Indoor / Outdoor or Urban / Rural / Forest
Which object classes?

Key Papers

- Point cloud segmentation
  - PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation
    https://github.com/charlesq34/pointnet
    *Both used in the two following papers
  - PointNet++: deep hierarchical feature learning on point sets in a metric space
    https://github.com/charlesq34/pointnet2
    *Rank 4 in http://www.semantic3d.net/
  - Large-scale point cloud segmentation with superpoint graphs
    https://github.com/loicland/superpoint_graph
    *Rank 1 in http://www.semantic3d.net/

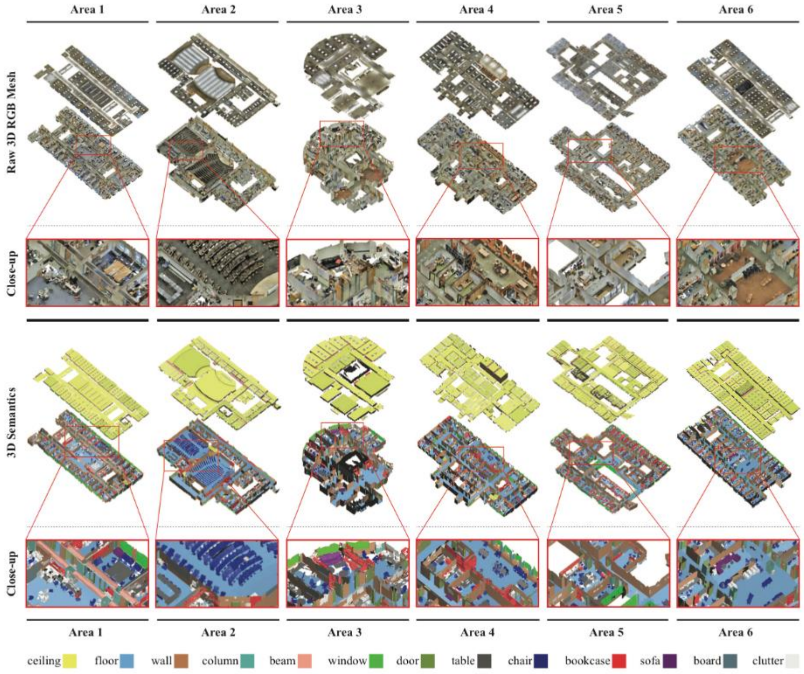
Publicly Available Dataset (1/3)

- Semantic3D
  - LiDAR dataset with over 3 billion points from a variety of urban and rural scenes.
  - http://www.semantic3d.net/
  - Managed by ETH (http://www.prs.igp.ethz.ch/)
  - 8 class labels, namely {1: man-made terrain, 2: natural terrain, 3: high vegetation, 4: low vegetation, 5: buildings, 6: hard scape, 7: scanning artefacts, 8: cars}.



Publicly Available Dataset (2/3)

- S3DIS (Stanford Large-Scale 3D Indoor Space)
  - 3D RGB point clouds of six floors from three different buildings
  - http://buildingparser.stanford.edu/dataset.html
  - Currently, *2D-3D-S dataset* is newly released.
  - 13 object classes (ceiling, floor, wall, beam, column, window, door, and movable elements: table, chair, sofa, bookcase, board and clutter for all other elements)

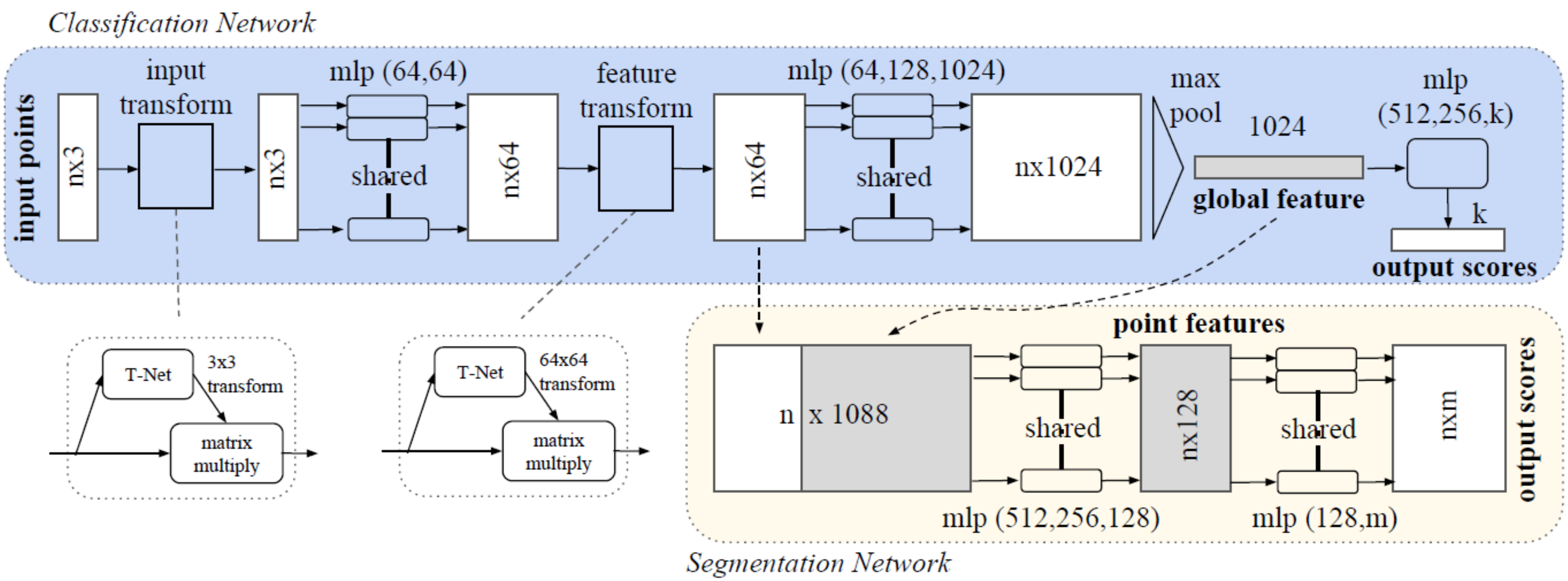


Publicly Available Dataset (3/3)

- Etc
  - Oakland 3-D Point Cloud Dataset (2009)
    - http://www.cs.cmu.edu/~vmr/datasets/oakland_3d/cvpr09/doc/
  - NYU Depth Dataset V2 (2012)
    - https://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html
  - Sydney Urban Objects data set
    - http://www.acfr.usyd.edu.au/papers/SydneyUrbanObjectsDataset.shtml
  - IQmulus & TerraMobilita Contest
    - Mobile laser scans (MLS) in dense urban environments
    - http://data.ign.fr/benchmarks/UrbanAnalysis/
  - Vaihingen3D airborne benchmark
    - http://www2.isprs.org/commissions/comm3/wg4/3d-semantic-labeling.html

PointNet

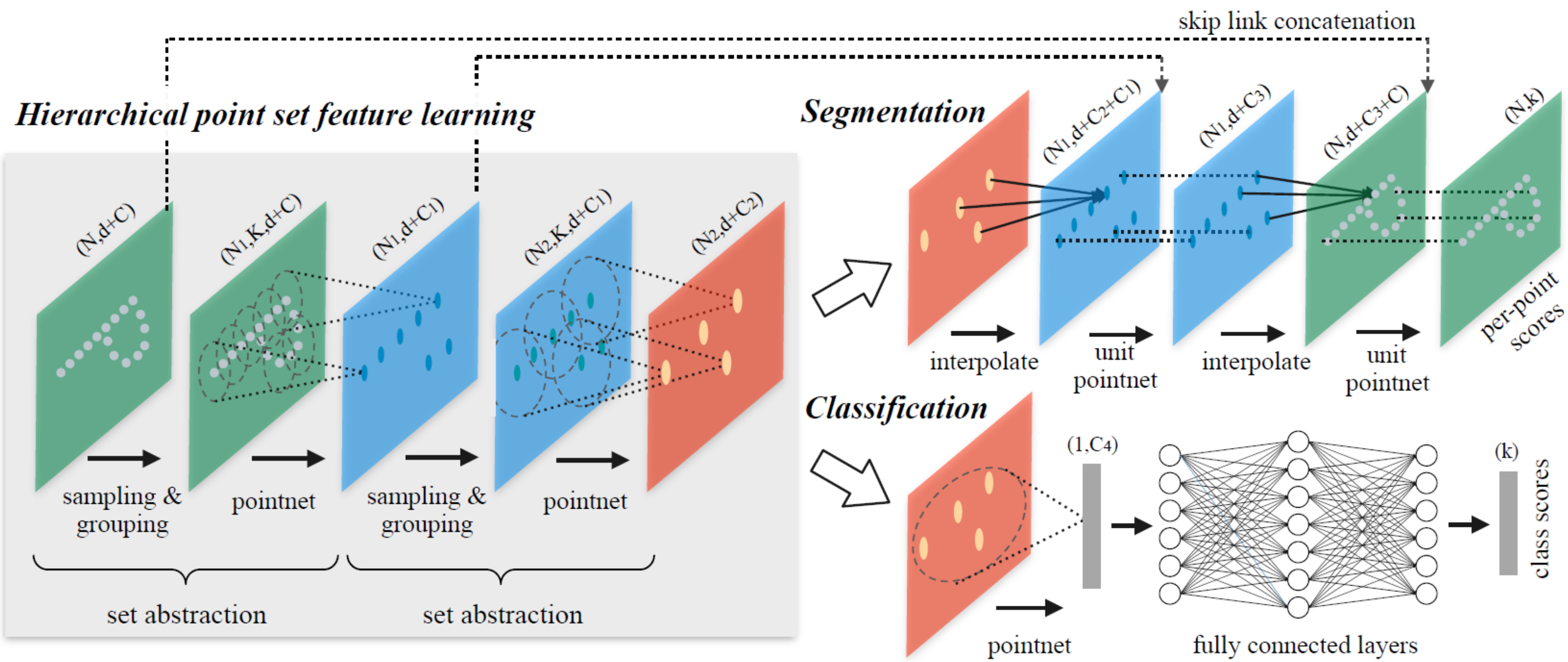
- Architecture
  - http://stanford.edu/~rqi/pointnet/



CVPR 2017

PointNet++

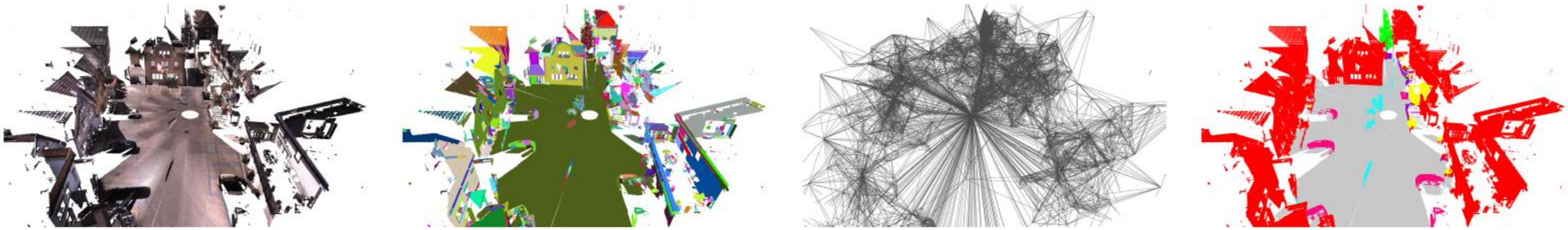
- Hierarchical Feature Learning Architecture
  - http://stanford.edu/~rqi/pointnet2/



NIPS 2017

Superpoint Graph (1/2)

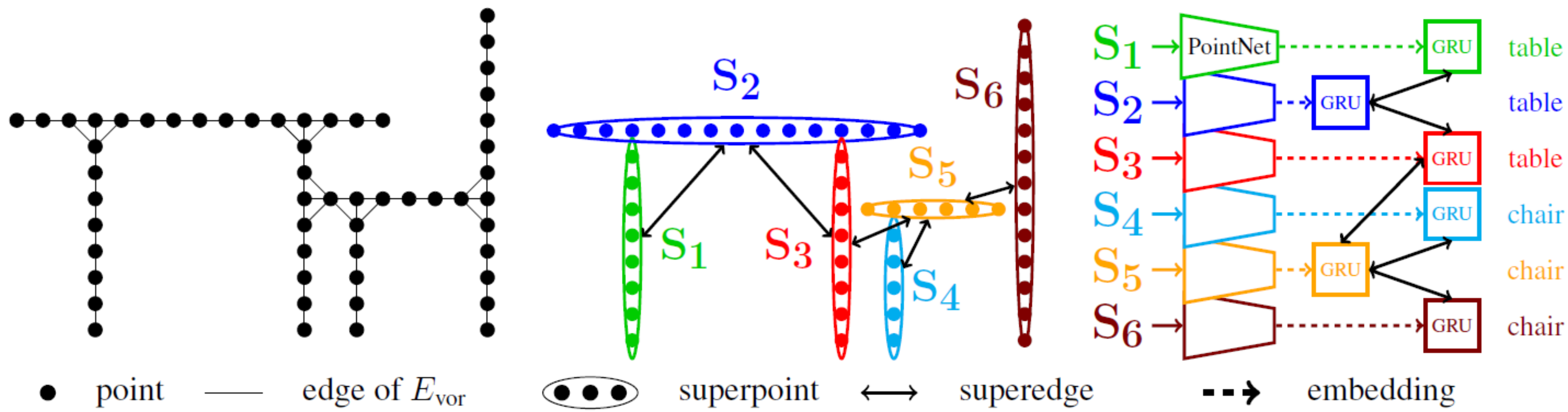
- Individual steps in pipeline



(a) RGB point cloud (b) Geometric partition (c) Superpoint graph (d) Semantic segmentation

CVPR 2018

Superpoint Graph (2/2)



(a) Input point cloud (b) Superpoint graph (c) Network architecture

Illustration of our framework on a toy scan of a table and a chair. We perform geometric partitioning on the point cloud (a), which allows us to build the superpoint graph (b). Each superpoint is embedded by a PointNet network. The embeddings are then refined in GRUs by message passing along superedges to produce the final labeling (c).

CVPR 2018

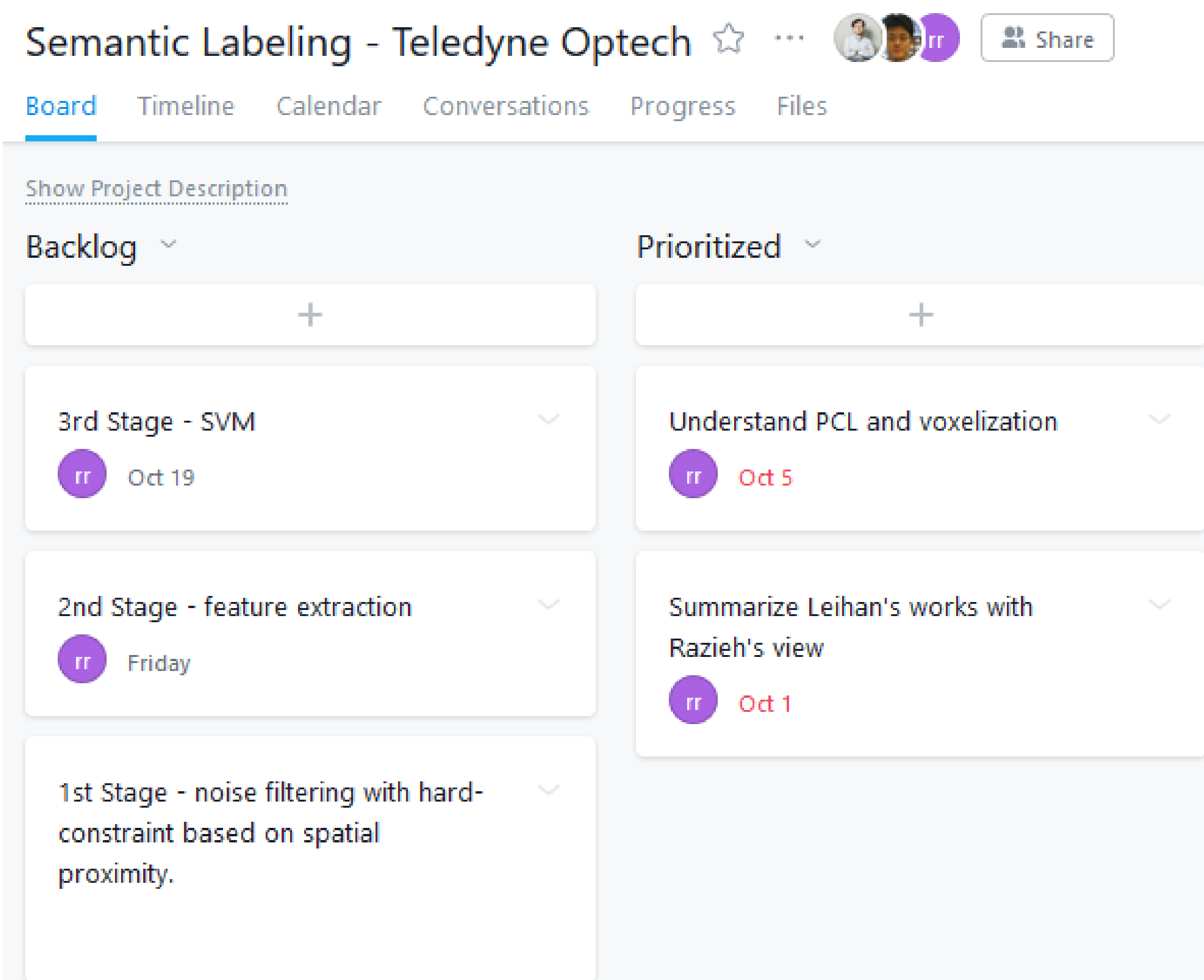
Plan

Month	Task	Deliverable
Oct 2018	• Problem definition • Dataset preparation • Literature survey	• Document describing problem definition, dataset, and literature survey • Visualization of dataset
Nov	• Practicing deep library • Design & implementation	• Document describing design
Dec	• Implementation	• Source code (Dec 31)
Jan 2019	• Documentation	• Document describing implementation (Jan 15)

Current Progress

Oct 19 2018

Asana Assignment



Notes

- Summarizing why noise happens, including some figures describing laser pulse and error band, and the terms like PIA and correspondence

- Knowing limitations of PointNet and how PointNet++ solve it
- Checking some parts (in Superpoint graph method) that the hand-crafted things are needed. (e.g. superpoint creation, superedge features)
- Knowing limitations of superpoint graph method
- Understanding how each method deal with object-scale issue and contextual information

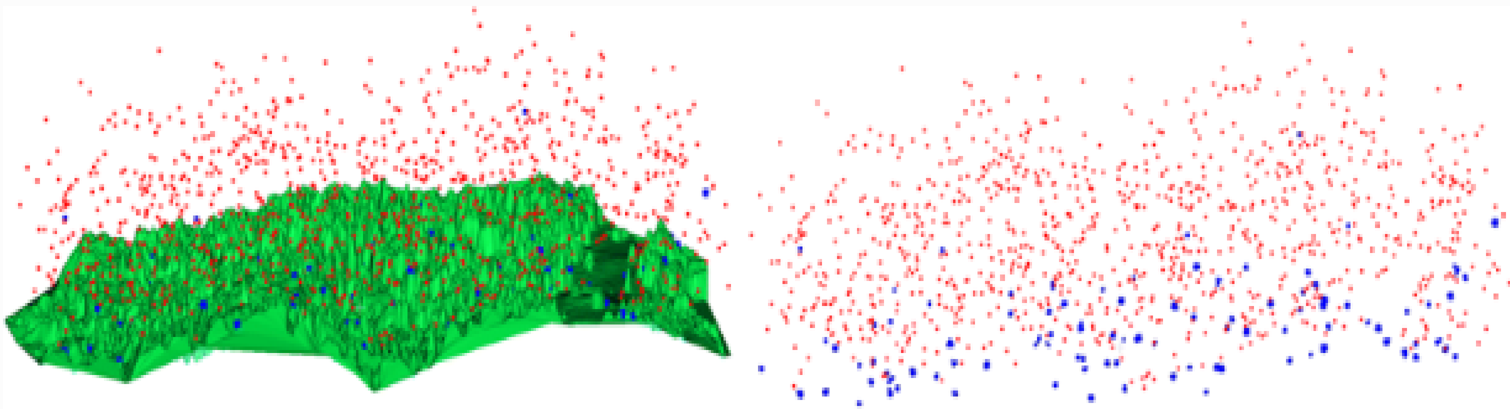
- Shenlong Wang's paper
  - Deep Parametric Continuous Convolutional Neural Networks

Noise filtering (1/2)

LiDAR is a relatively new technology which is an alternative to field surveying and photogrammetric techniques t o collect elevation data. This technology is able to provide high accuracy three-dimensional data with reasonable cost and time. 3-D data acquired by this technology are applicable in 3-D urban modeling, DTM generation, map ping and etc. Although LiDAR data present high height accuracy, there are some defects in them leading to some disadvantages in output of next processes. Hence, these errors should be removed before performing any proces s. One of the most important errors in LiDAR data is the outlier points. Measured elevation for these points is un reasonably more or less from their neighboring points. The outliers are mainly measurements that do not obey t he local surface geometry and do not belong to the topography of the interested area. In some references in the literature, the points with too high elevation values are named "positive outliers" and the points with too low el evation values are named "negative outliers", so we used these terms, too. The outliers can be caused from diffe rent sources. Positive outliers are resulted from hitting laser beams to suspended objects at high altitude like at mospheric articles. In connection with the negative outliers, it is believed that the laser beams be reflected sever al times among the glasses of buildings before they are detected, just like the multi-path effect of GPS. These sp ecular reflections result in a longer travel time of the laser beam, and thus a lower elevation is calculated during post-flight processing. The negative outliers are often located at a few spots beside which there are tall buildings

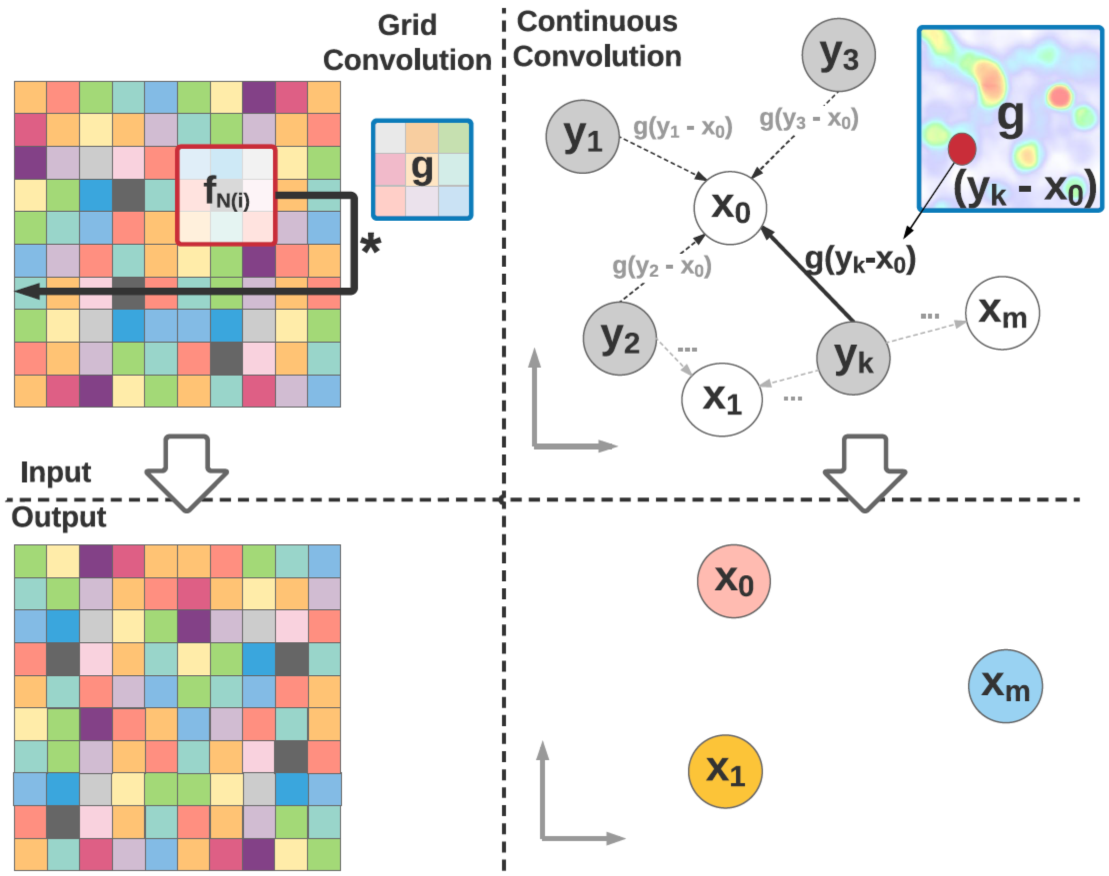
Noise filtering (2/2)

- Outliers in LiDAR data



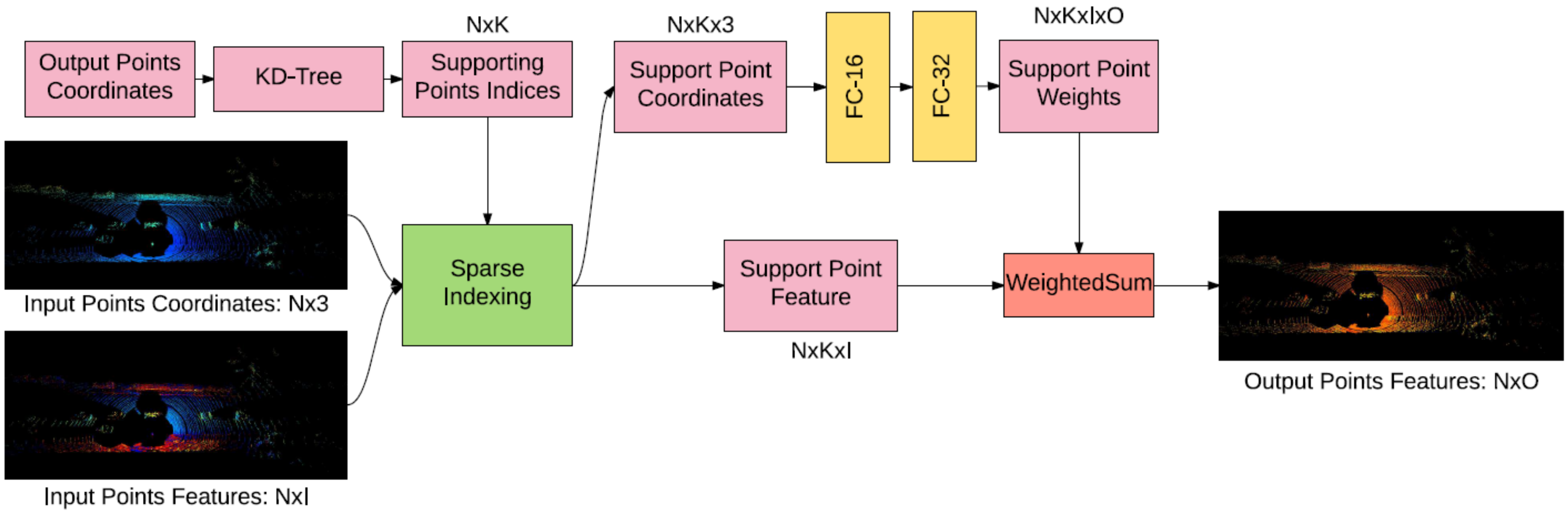
SfM dataset – left: clean DSM surface in green, random unclustered outliers in red, clustered random outliers in blue – right: only outliers

Deep Parametric Continuous CNNs (1/5)



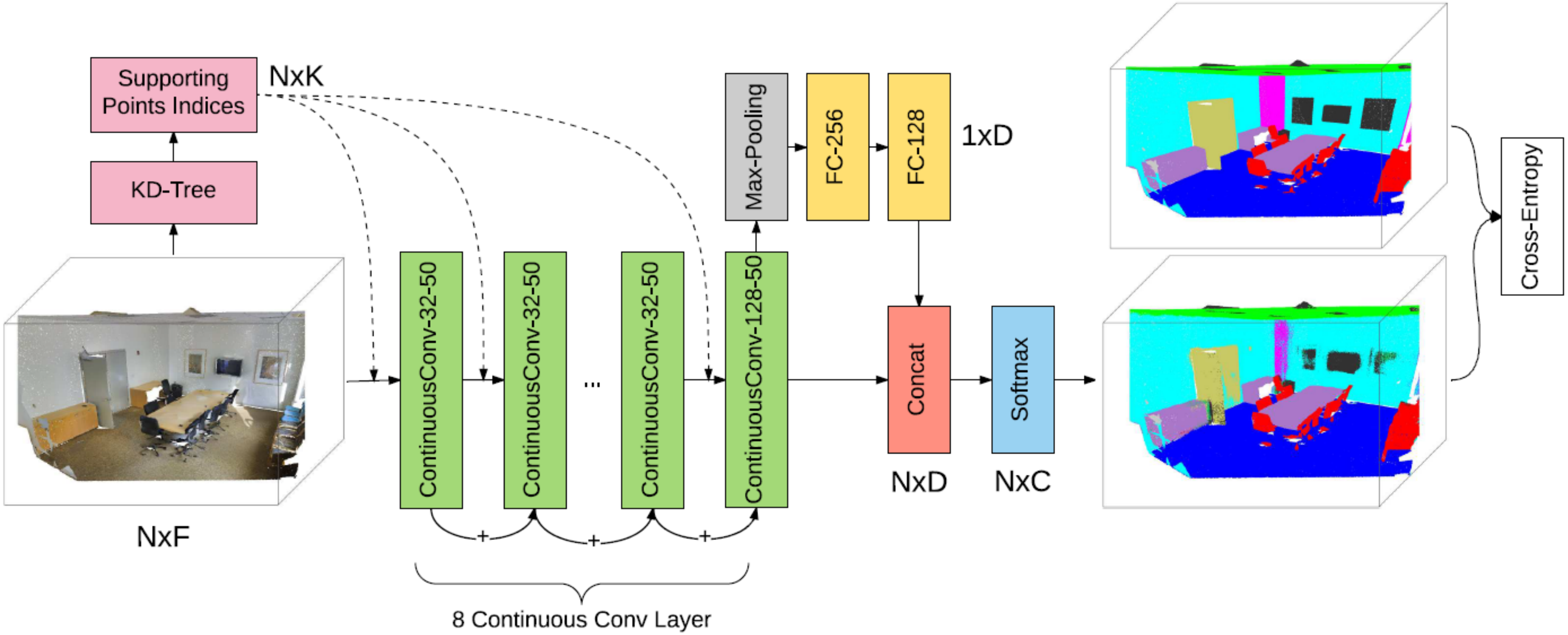
Unlike grid convolution, parametric continuous convolution uses kernel functions that are defined for arbitrary points in the continuous support domain.
As a result, it is possible to output features at points not seen in the input.

Deep Parametric Continuous CNNs (2/5)



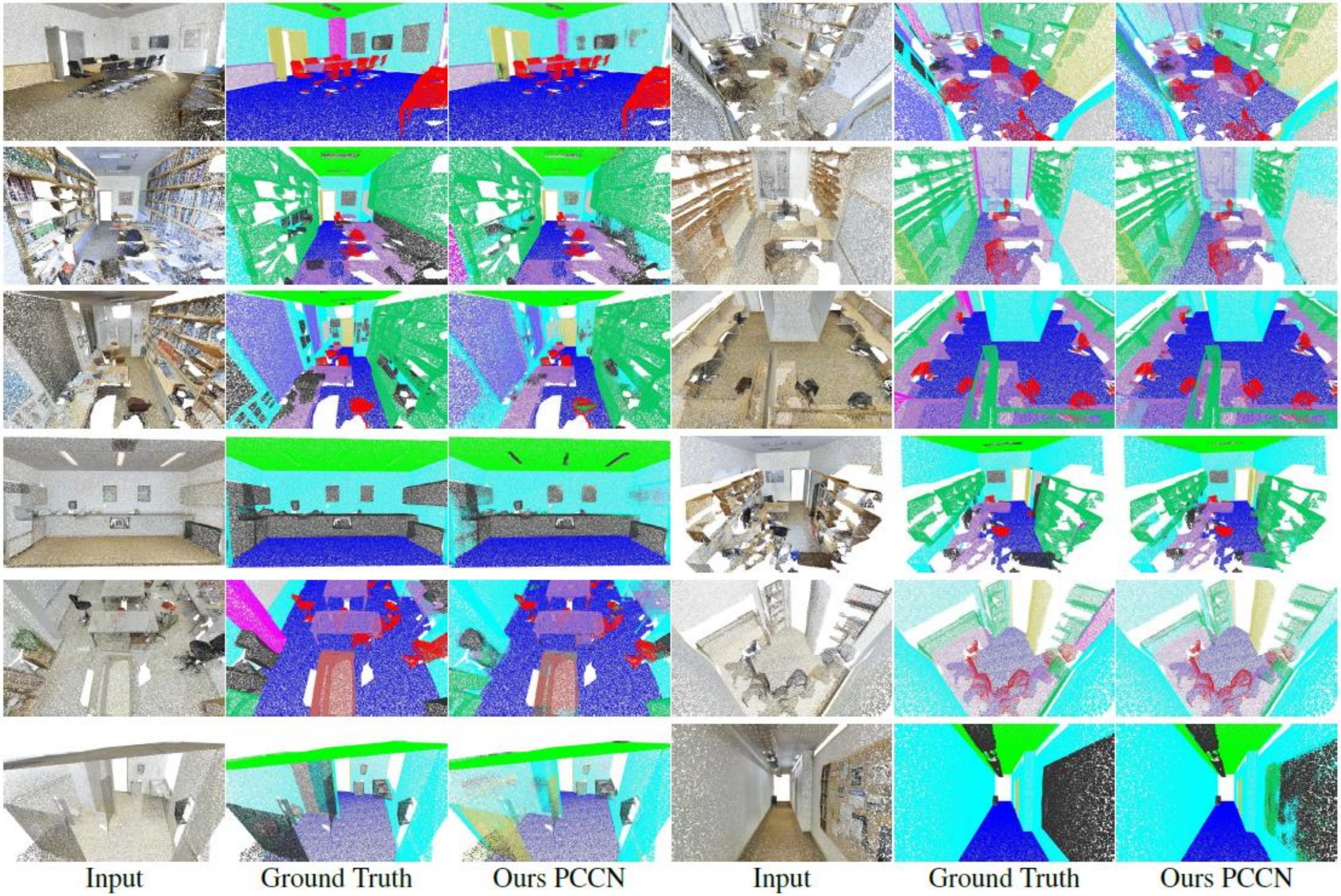
Detailed Computation Block for the Parametric Continuous Convolution Layer

Deep Parametric Continuous CNNs (3/5)



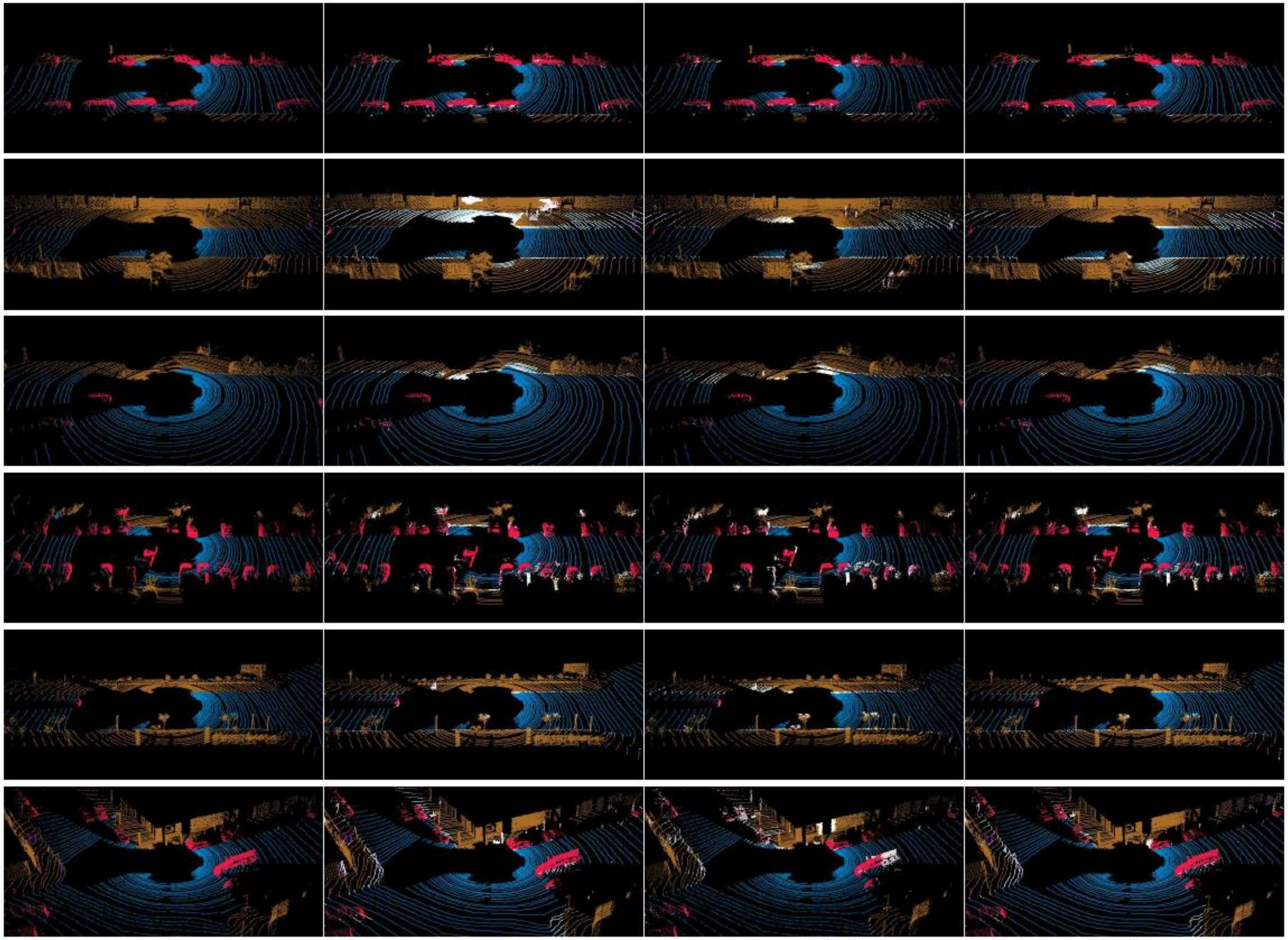
Architecture of the Deep Parametric Continuous CNNs for Semantic Labeling Task

Deep Parametric Continuous CNNs (4/5)



Semantic Segmentation Results on Stanford Indoor3D Dataset

Deep Parametric Continuous CNNs (5/5)



Semantic Segmentation Results on Driving Scene Dataset;
Colored: correct prediciton; white: wrong prediciton.

Plan

Month	Task	Deliverable
Oct 2018	• Problem definition • Dataset preparation • Literature survey	• Document describing problem definition, dataset, and literature survey • Visualization of dataset
Nov	• Practicing deep library • Design & implementation	• Document describing design
Dec	• Implementation	• Source code (Dec 31)
Jan 2019	• Documentation	• Document describing implementation (Jan 15)

Current Progress

Nov 5 2018

Objectives

- Building 1st ver deep network for each task

Task	Major Contributor	Objectives
Noise filtering	Razieh Ramak	Point cloud segmentation (Noise/Non-noise), non real-time
Point cloud segmentation	Maryam Jameela	Point cloud segmentation (N-class objects), non real-time
3D object detection	Jungwon Kang	Real-time 3D object detection

Plan

Month	Task	Deliverable
Oct 2018	• Problem definition • Dataset preparation • Literature survey	• Document describing problem definition, dataset, and literature survey • Visualization of dataset
Nov	• Practicing deep library • Design & implementation	• Document describing design
Dec	• Implementation	• Source code (Dec 31)
Jan 2019	• Documentation	• Document describing implementation (Jan 15)

*Submission deadline of major conferences starts from March.

Paper List

- Papers we have read (for point cloud segmentation)

Title	Conference
PointNet: deep learning on point sets for 3D classification and segmentation	CVPR 2017
PointNet++: deep hierarchical feature learning on point sets in a metric space	NIPS 2017
Large-scale point cloud segmentation with superpoint graphs	CVPR 2018
Deep parametric continuous convolutional neural networks	CVPR 2018
SPLATNet: sparse lattice networks for point cloud processing	CVPR 2018

SPLATNet:Sparse Lattice Networks (1/2)

- SPLATNet

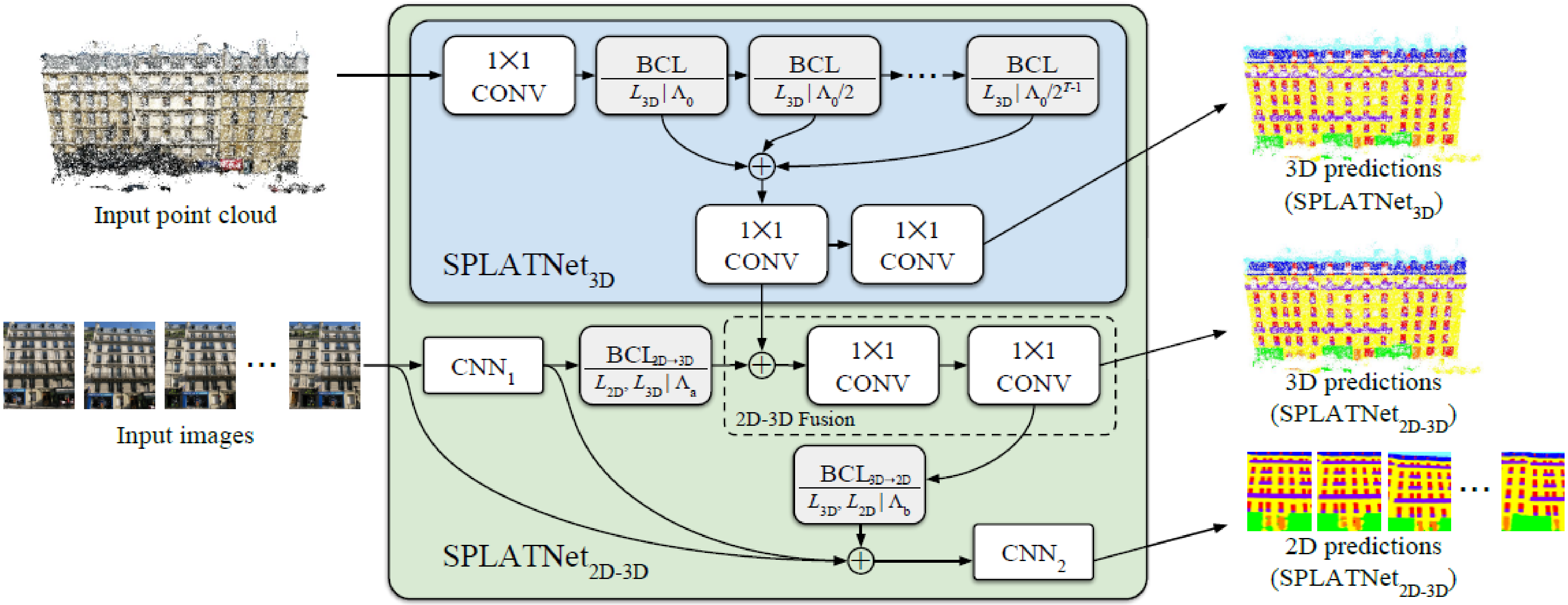
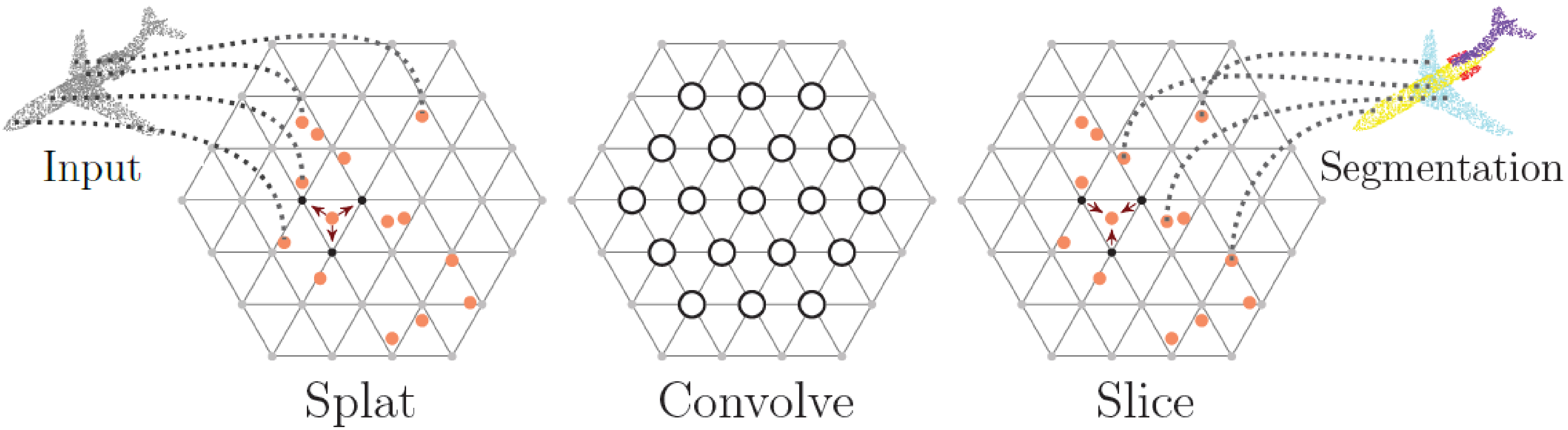


Illustration of inputs, outputs and network architectures for SPLATNet$_{3D}$ and SPLATNet$_{2D\text{-}3D}$.

SPLATNet:Sparse Lattice Networks (2/2)

- Bilateral Convolution Layer



Splat: BCL first interpolates input features F onto a d_l-dimensional permutohedral lattice defined by the lattice features L at input points.

Convolve: BCL then does d_l-dimensional convolution over this sparsely populated lattice.

Slice: The filtered signal is then interpolated back onto the input signal. For illustration, input and output are shown as point cloud and the corresponding segmentation labels.

What to do (during Nov)

- Continuing literature survey
  - Thinking of what the fundamental issues are unresolved
  - Writing document about the literatures
- Practicing deep library
- Initial design of our own network

Current Progress

Nov 15 2018

Contents

- Razieh's Progress on Noise Filtering
- Paper Review
- Future Plan

Paper Review

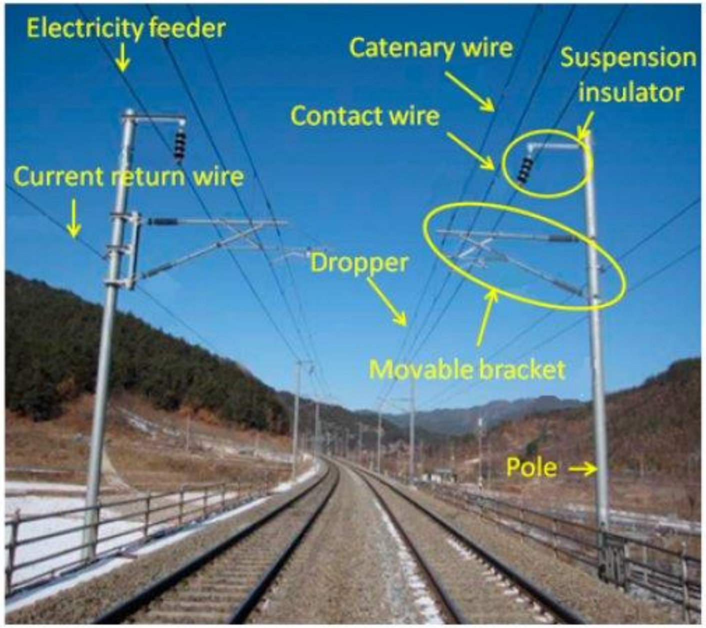
Multi-Range CRF (1/5)

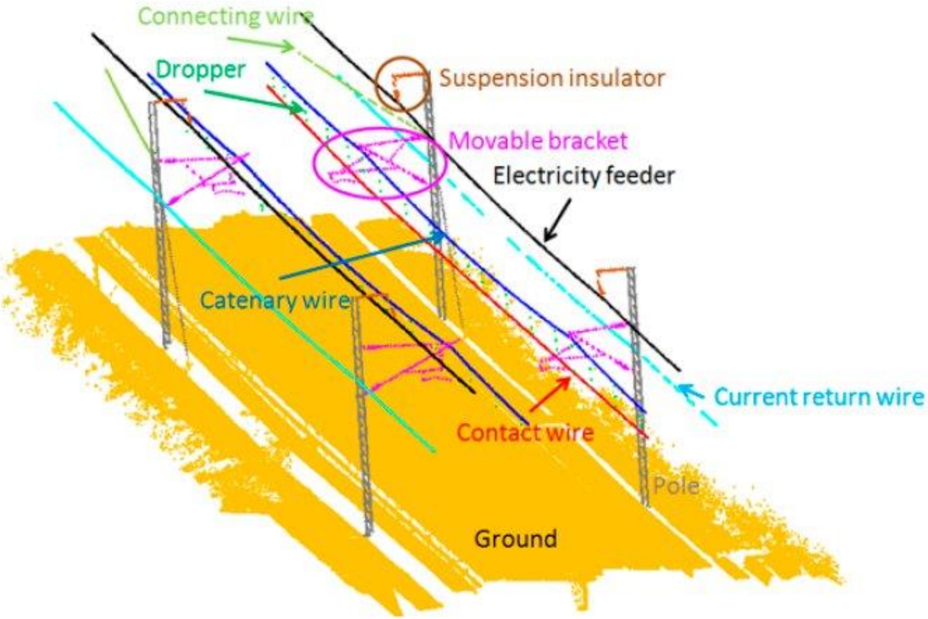
Multi-Range CRF for Classifying Railway Electrification System Objects Using Mobile Laser Scanning Data, J. Jung, et al., Remote Sensing 2016

Trimble MX8 mounted on a train

Parameter	Values
Accuracy	10 mm
Precision	5 mm
Maximum effective measurement rate	600,000 points/second
Line scan speed	Up to 200 lines/second
Echo signal intensity	High resolution 16-bit intensity
Range	Up to 500 m

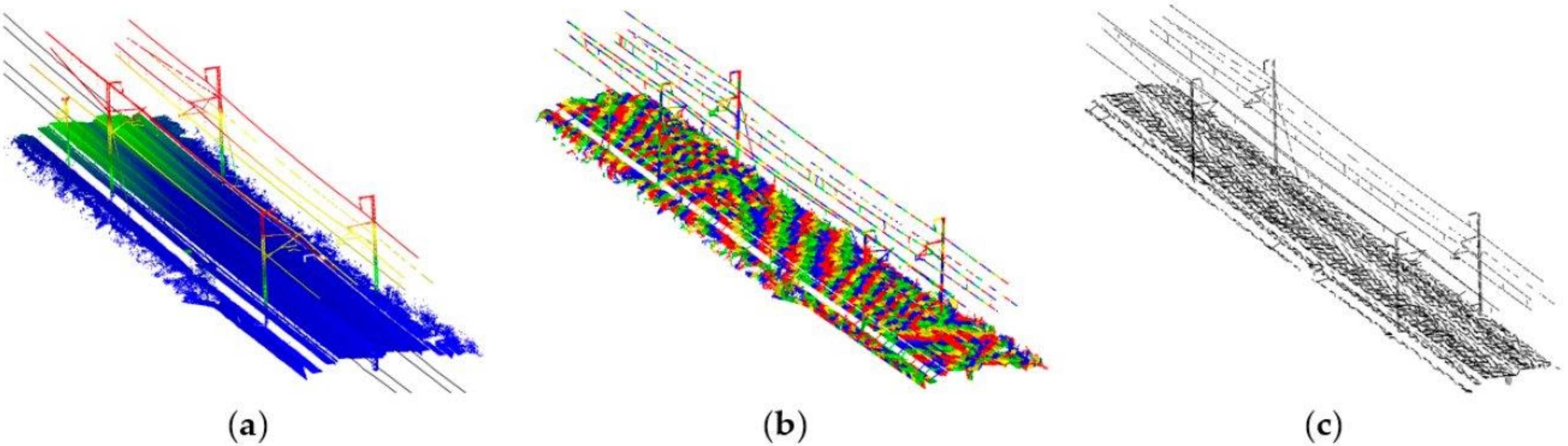
Multi-Range CRF (2/5)





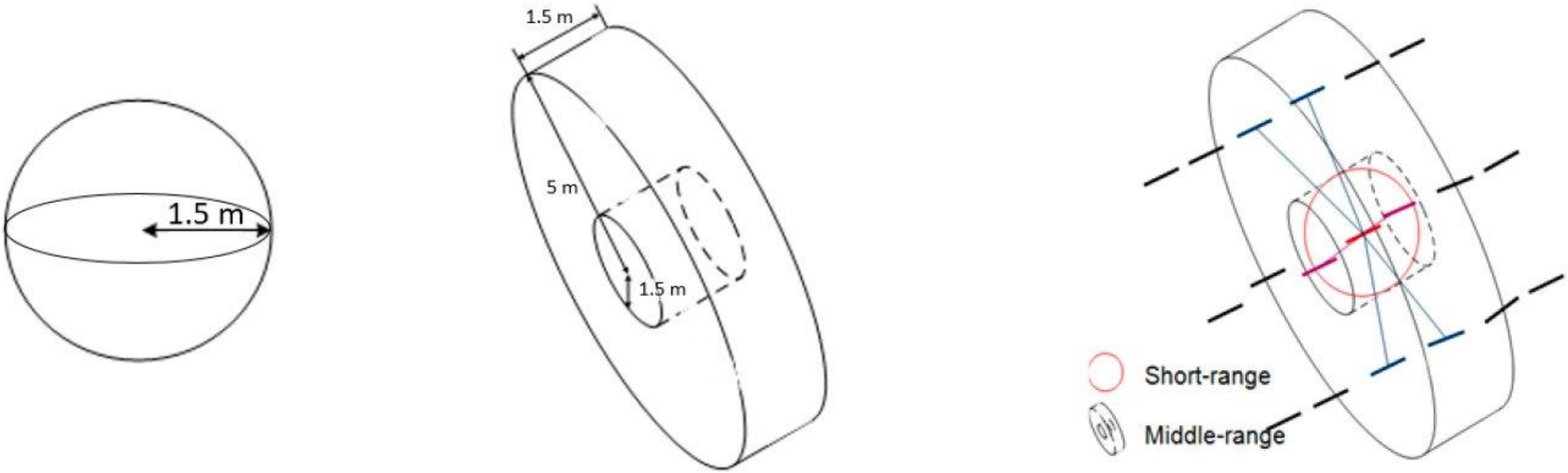
Electrification system configuration and 10 object classes of the railway
a photograph (left), and Mobile Laser Scanning (MLS) data (right)

Multi-Range CRF (3/5)



Example of voxelization and line extraction:
(a) input MLS data and rail vectors, (b) voxelization, and (c) extracted lines

Multi-Range CRF (4/5)



(a) (b) (c)

Neighboring systems: (a) for short-range graph, (b) for long-range graph, and (c) combined neighboring systems

Multi-Range CRF (5/5)

- Summary
  - Handled line segments rather than raw point clouds
  - (Short & Long) Multi-range CRF (which is a graphical model)

  - The 'Superpoint Graph' paper has a similar algorithm structure.

Future Plan

Practicing Deep Library

- Producing initial results, while practicing deep library
  - Within this year (about 40 days remained)
  - Task 1: point segmentation (for Optech) on Semantic3D dataset
    - ex: PointNet
  - Task 2: object or lane detection (for Thales-AVIN)

Risk: only one available workstation

Further Reading (1/2)

- Derek Hoiem
  - http://dhoiem.cs.illinois.edu/

- Hoiem's papers (sequential segmentation-related)
  - 3D-PRNN: Generating Shape Primitives with Recurrent Neural Networks, ICCV 2017
  - Learning to Localize Little Landmarks, CVPR 2016

Further Reading (2/2)

sequential segmentation-related

- ReSeg: A Recurrent Neural Network-based Model for Semantic Segmentation
  - https://arxiv.org/pdf/1511.07053.pdf
- Recurrent Segmentation for Variable Computational Budgets
  - https://arxiv.org/pdf/1711.10151.pdf
- Combined convolutional and recurrent neural networks for hierarchical classification of images
  - https://arxiv.org/pdf/1809.09574.pdf
- CNN-RNN: A Unified Framework for Multi-label Image Classification
  - https://arxiv.org/ftp/arxiv/papers/1604/1604.04573.pdf

Current Progress

Nov 28 2018

Contents

- Razieh's Progress on Noise Filtering
- Paper Review
- Future Plan

Paper Review

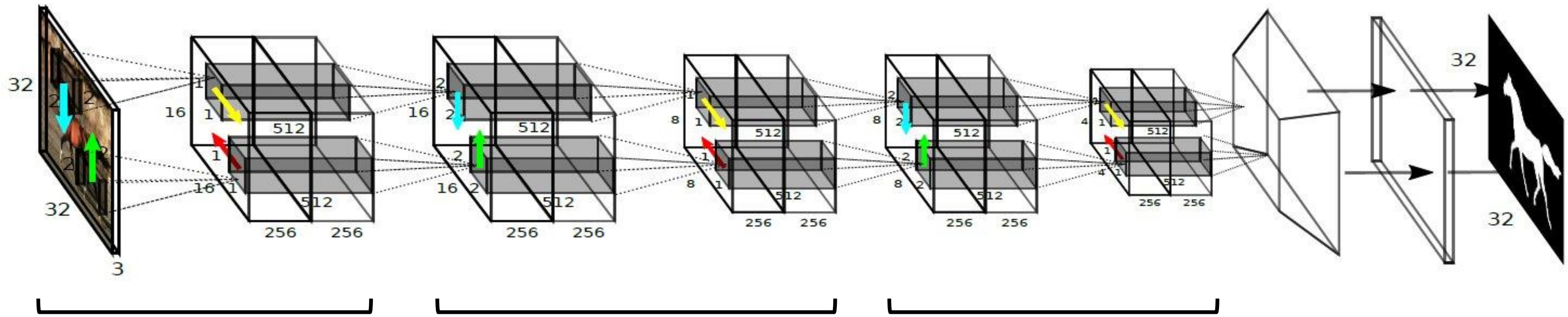
sequential segmentation-related

- ReSeg: A Recurrent Neural Network-based Model for Semantic Segmentation
  - https://arxiv.org/pdf/1511.07053.pdf

- 3D Recurrent Neural Networks with Context Fusion for Point Cloud Semantic Segmentation
  - http://openaccess.thecvf.com/content_ECCV_2018/papers/Xiaoqing_Ye_3D_Recurrent_Neural_ECCV_2018_paper.pdf

ReSeg (1/2)

ReSeg: A Recurrent Neural Network-based Model for Semantic Segmentation



The ReSeg network.

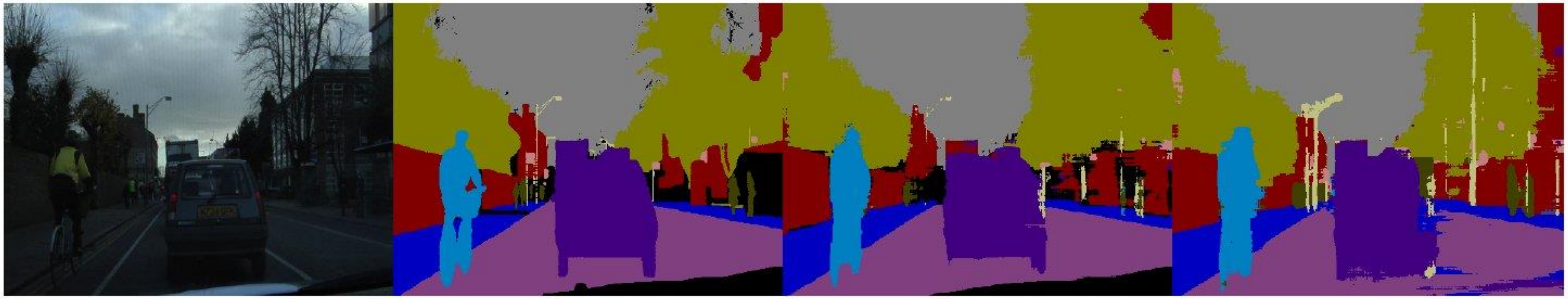
(※ The input to ReSeg is preprocessed by the pretrained VGG-16 convolutional layer, although the layer is shown here.)

The first 2 RNNs (blue and green) are applied on 2x2x3 patches of the image, their 16x16x256 feature maps are concatenated and fed as input to the next two RNNs (red and yellow) which read 1x1x512 patches and emit the output of the first ReNet layer.

Two similar ReNet layers are stacked, followed by an upsampling layer and a softmax nonlinearity.

ReSeg (2/2)

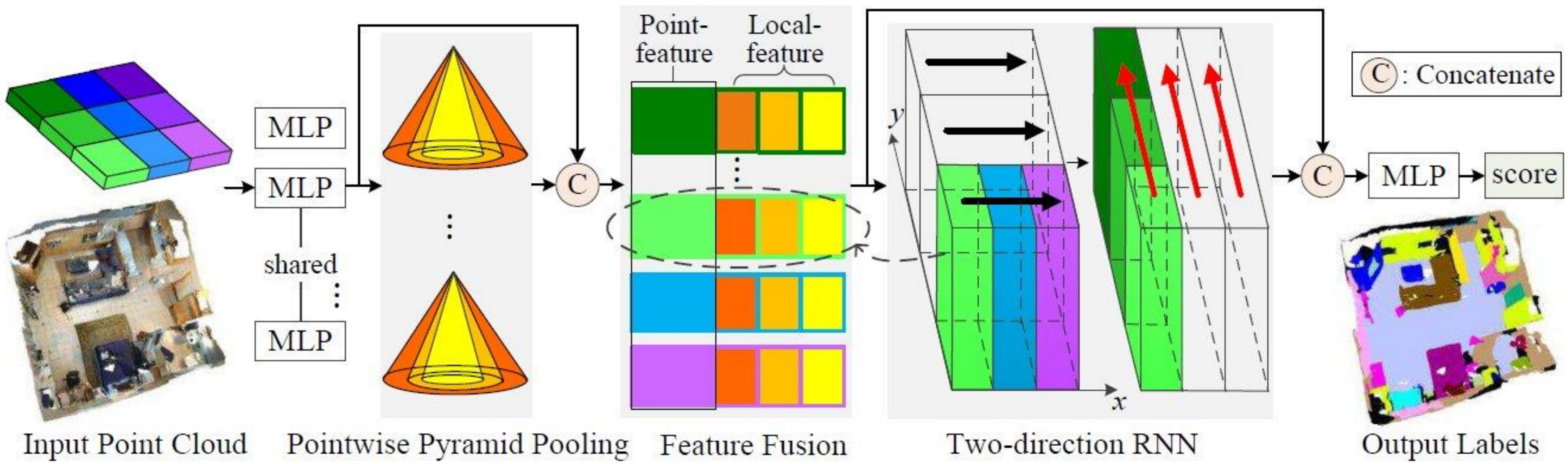
ReSeg: A Recurrent Neural Network-based Model for Semantic Segmentation



Input image | Ground truth | ReSeg | ReSeg (with class balancing)

3D RNN (1/4)

3D Recurrent Neural Networks with Context Fusion for Point Cloud Semantic Segmentation

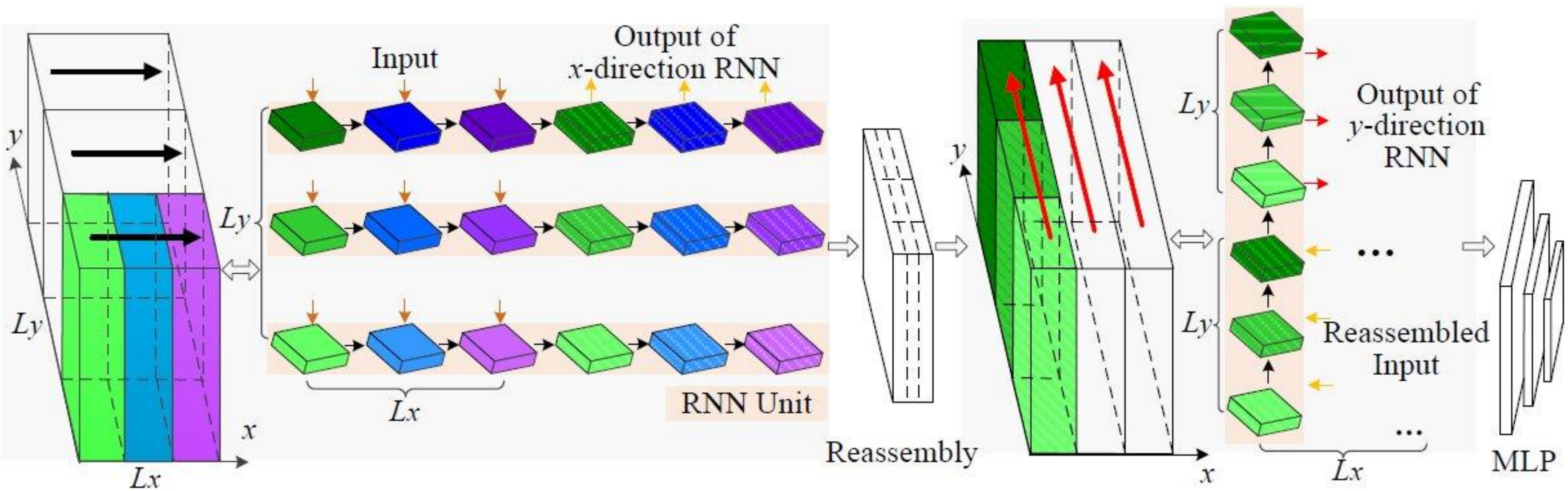


Overview of the proposed approach

The architecture takes as input the unstructured point cloud and outputs pointwise semantic labels. Point-features and local cell features are concatenated and passed through the two-direction RNN module along x and y.
The output of the first RNNs (black arrowed) are reorganized and fed to the next RNNs (red).

3D RNN (2/4)

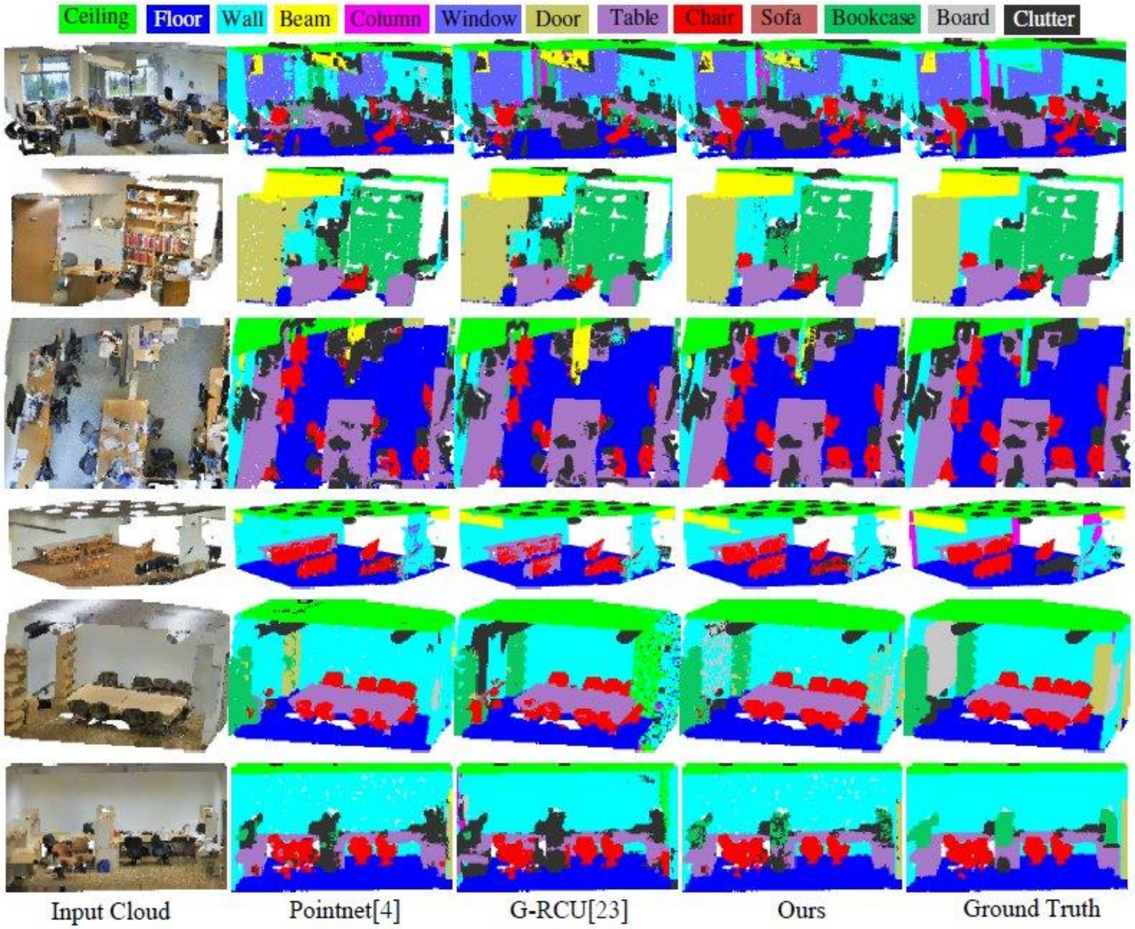
3D Recurrent Neural Networks with Context Fusion for Point Cloud Semantic Segmentation



Overview of the proposed approach

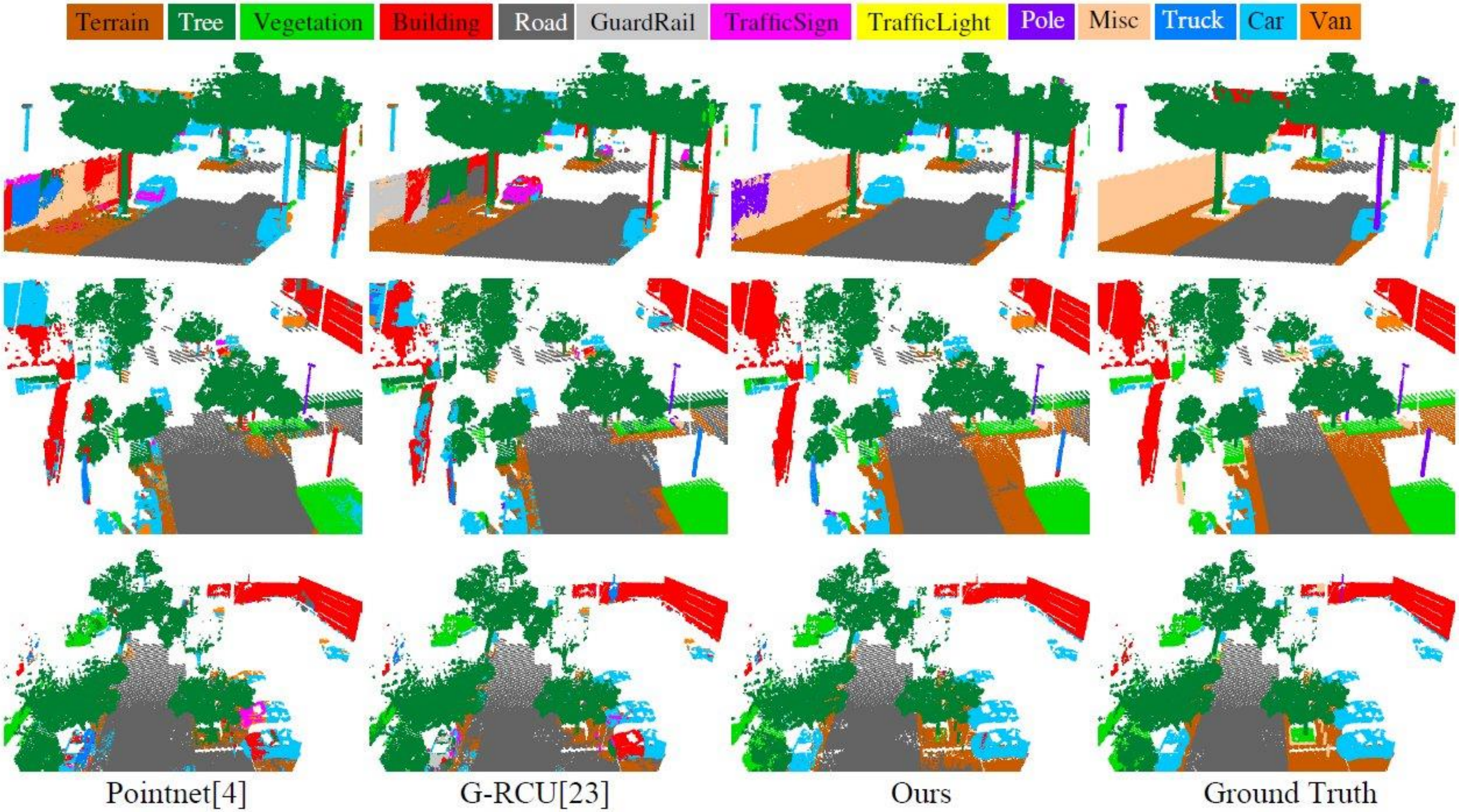
The architecture takes as input the unstructured point cloud and outputs pointwise semantic labels. Point-features and local cell features are concatenated and passed through the two-direction RNN module along x and y.
The output of the first RNNs (black arrowed) are reorganized and fed to the next RNNs (red).

3D RNN (3/4)



Results on indoor S3DIS dataset

3D RNN (4/4)



Results on outdoor virtual KITTI dataset

Future Plan (for Point Segmentation)

Dec 2018	• Understanding dataset & visualize it Optech: Mobile Lynx Data (Sept 2018), Titan and Galaxy (Oct 2018) Public: Semantic3D, S3DIS, KITTI • Practicing deep library (without the use of GPU) • Literature review
Jan 2019	• Applying an existing network (e.g. PointNet) to the Optech dataset • Literature review

- Also needs for a plan for noise filtering, object detection
- Jungwon's workstation will be fully used by at least Dec 11 (for other project).