Classification of Point Cloud for Road Scene Understanding with Multiscale Voxel Deep Network

Xavier Roynard

Jean-Emmanuel Deschaud

François Goulette

xavier.roynard@mines-paristech.fr, jean-emmanuel.deschaud@mines-paristech.fr, francois.goulette@mines-paristech.fr

October 1, 2018

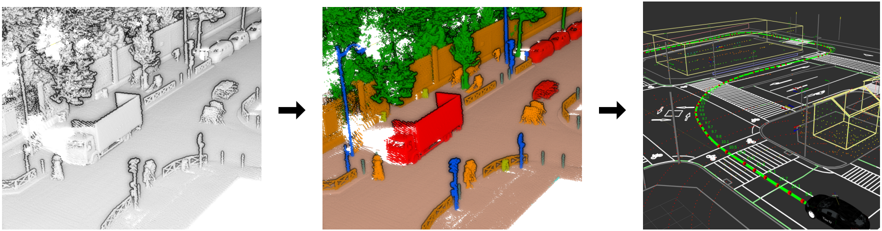
Presentation Outline

Context

Autonomous vehicles require HD-Maps for navigation and decision-making process

A production pipeline of HD-Maps can be :

- 3D point cloud acquisition by Mobile Laser Scanning (MLS),
- Precise 3D localization of relevant objects (road signs and ground markings),
- Extraction of mobile objects,
- Detection of navigation area and buildings.



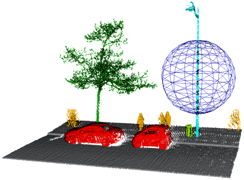
State of the Art

Point-wise Classification

- Hand-Made Features (dimensionality attributes, multi-scale) [a],
- Deep Learning on Voxel Grid Neighborhood [b]

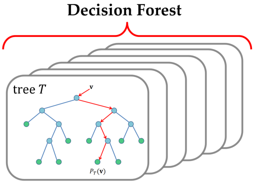
a. Timo Hackel, Jan D Wegner et Konrad Schindler. “Fast semantic segmentation of 3D point clouds with strongly varying density”. In : *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Prague, Czech Republic* 3 (2016), p. 177–184.

b. Jing Huang et Suya You. “Point cloud labeling using 3d convolutional neural network”. In : *Pattern Recognition (ICPR), 2016 23rd International Conference on*. IEEE. 2016, p. 2670–2675.



$$\rightarrow \quad X = \begin{bmatrix} X_1 \\ X_2 \\ \\ X_d \end{bmatrix} \quad \rightarrow$$

Feature vector



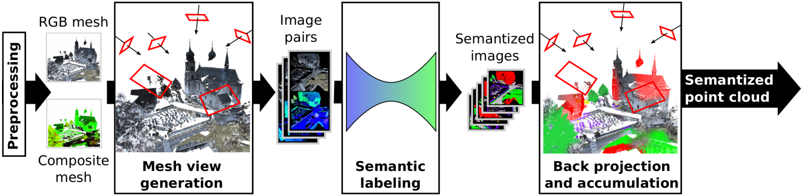
State of the Art

Region-wise Classification

- on images : Snapnet_ [a]
- on voxel Grid : SEGCloud [b]

a. Alexandre Boulch, Bertrand Le Saux et Nicolas Audebert. "Unstructured point cloud semantic labeling using deep segmentation networks". In : *Eurographics Workshop on 3D Object Retrieval*. T. 2. 2017, p. 1.

b. Lyne P Tchapmi et al. "SEGCloud : Semantic Segmentation of 3D Point Clouds". In : *arXiv preprint arXiv :1710.07563* (2017).



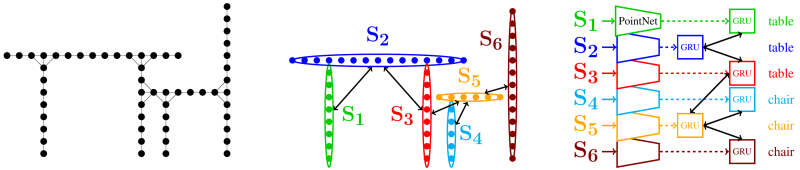
State of the Art

Segmentation-based Classification

- SPGraph [a]

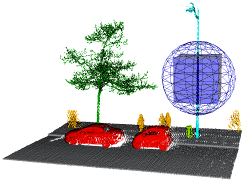
---

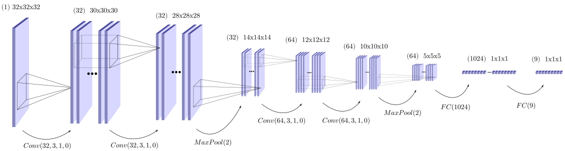
a. Loic LANDRIEU et Martin SIMONOVSKY. "Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs". In : *arXiv preprint arXiv :1711.09869* (nov. 2017).



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

Our Approach





Training a Deep Neural Network on fully annotated 3D point cloud scenes

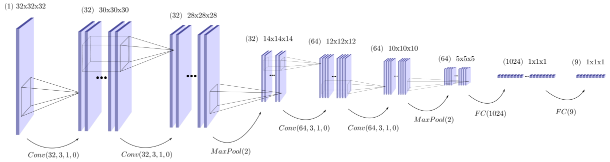
- Some challenges :
  - very unbalanced classes,
  - most represented classes are also the least geometrically diversified (groud, buildings),
  - billion of samples.
- Using all samples (points) in one epoch would be infeasible.

Proposed solution

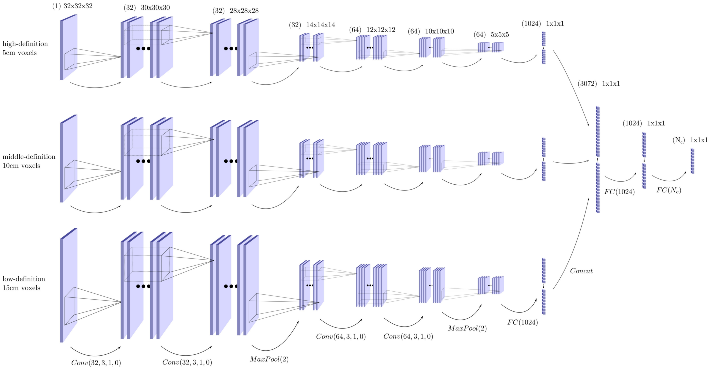
- randomly sample $N > 0$ points in each class of the training dataset,
- then one epoch is : pass randomly all sampled points in the network

Multi-Scale Architecture

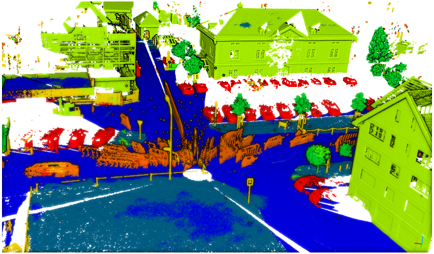
Mono-Scale

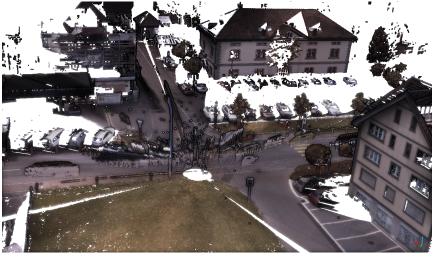


Multi-Scale



Results on Semantic3D





Rank	Method	Averaged IoU	Overall Accuracy	Per class IoU: man-made terrain	natural terrain	high vegetation	low vegetation	buildings	hard scape	scanning artefacts	cars
1	SPGraph [1]	**73.2%**	**94.0%**	**97.4%**	**92.6%**	**87.9%**	**44.0%**	**93.2%**	31.0%	**63.5%**	76.2%
2	MS3_DVS**(Ours)**	65.3%	88.4%	83.0%	67.2%	83.8%	36.7%	92.4%	**31.3%**	50.0%	**78.2%**
3	RF_MSSF [2]	62.7%	90.3%	87.6%	80.3%	81.8%	36.4%	92.2%	24.1%	42.6%	56.6%
4	SegCloud [3]	61.3%	88.1%	83.9%	66.0%	86.0%	40.5%	91.1%	30.9%	27.5%	64.3%
5	SnapNet_ [4]	59.1%	88.6%	82.0%	77.3%	79.7%	22.9%	91.1%	18.4%	37.3%	64.4%

1. Loic Landrieu et Martin Simonovsky. "Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs". In : *arXiv preprint arXiv :1711.09869* (nov. 2017).

2. Hugues Thomas et al. "Semantic Classification of 3D Point Clouds with Multiscale Spherical Neighborhoods". In : *arXiv preprint arXiv :1808.00495* (2018).

3. Lyne P Tchapmi et al. "SEGCloud : Semantic Segmentation of 3D Point Clouds". In : *arXiv preprint arXiv :1710.07563* (2017).

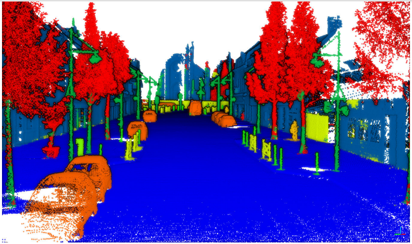
4. Alexandre Boulch, Bertrand Le Saux et Nicolas Audebert. "Unstructured point cloud semantic labeling using deep segmentation networks". In : *Eurographics Workshop on 3D Object Retrieval*. T. 2. 2017, p. 1.

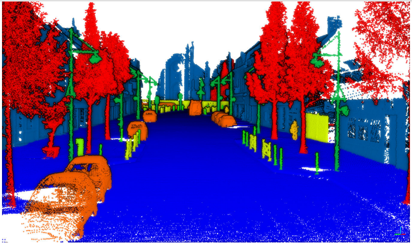
Results on Paris-Lille-3D

New Benchmark for Point Cloud Classification : Paris-Lille-3D [a] :

- Training set : 140 million manually annotated points, 50 classes, $2\mathrm{km}$, 2 cities
- Test set : 30 million points, 9 classes, 2 other cities

a. X. Roynard, J.-E. Deschaud et F. Goulette. "Paris-Lille-3D : a large and high-quality ground truth urban point cloud dataset for automatic segmentation and classification". In : *ArXiv e-prints* (nov. 2017). arXiv : 1712.00032 [cs.LG].





Rank	Method	Averaged IoU	Per class IoU: ground	building	pole	bollard	trash can	barrier	pedestrian	car	natural
1	MS3_DVS**(Ours)**	**66.89%**	99.03%	**94.76%**	**52.40%**	38.13%	**36.02%**	**49.27%**	**52.56%**	**91.3%**	**88.58%**
2	RF_MSSF [5]	56.28%	**99.25%**	88.63%	47.75%	**67.27%**	2.31%	27.09%	20.61%	74.79%	78.83%

5. Hugues Thomas et al. "Semantic Classification of 3D Point Clouds with Multiscale Spherical Neighborhoods". In : *arXiv preprint arXiv :1808.00495* (2018).

Comparison Mono/Multi-scales

Precision and Recall on Paris-Lille-3D

Improvement on most of the classes.

Class	Precision		Recall	
	MS3_DVS	MS1_DVS	MS3_DVS	MS1_DVS
ground	**97.74%**	97.08%	**98.70%**	98.28%
buildings	**85.50%**	84.28%	**95.27%**	90.65%
poles	**93.30%**	92.27%	92.69%	**94.16%**
bollards	98.60%	**98.61%**	93.93%	**94.16%**
trash cans	**95.31%**	93.52%	79.60%	**80.91%**
barriers	**85.70%**	81.56%	**77.08%**	73.85%
pedestrians	**98.53%**	93.62%	**95.42%**	92.89%
cars	93.51%	**96.41%**	**98.38%**	97.71%
natural	**89.51%**	88.23%	**92.52%**	91.53%

Mean F1 Score

The contribution of multi-scale network is obvious.

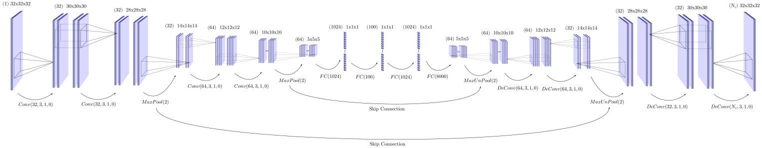
Dataset \ Method	MS3_DVS	MS1_DVS	VoxNet [6]
Paris-Lille-3D	**89.29%**	88.23%	86.59%
Semantic3D	**79.36%**	74.05%	71.66%

6. Daniel Maturana et Sebastian Scherer. "VoxNet : A 3D convolutional neural network for real-time object recognition". In : *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE. 2015, p. 922–928

Work in progress

Work in progress

- Use network architectures closer to the state of the art (Inception/ResNet).
- Adapt the Multi-Scale architecture to U-Net networks for semantic segmentation.
- Get closer to real-time inference with an Octree structure.
- Ensemble on several networks or several orientations of input point cloud.



Thank you !

Questions ?

Bibliographie

Alexandre Boulch, Bertrand Le Saux et Nicolas Audebert. “Unstructured point cloud semantic labeling using deep segmentation networks”. In : *Eurographics Workshop on 3D Object Retrieval*. T. 2. 2017, p. 1.

Timo Hackel, Jan D Wegner et Konrad Schindler. “Fast semantic segmentation of 3D point clouds with strongly varying density”. In : *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Prague, Czech Republic* 3 (2016), p. 177–184.

Jing Huang et Suya You. “Point cloud labeling using 3d convolutional neural network”. In : *Pattern Recognition (ICPR), 2016 23rd International Conference on*. IEEE. 2016, p. 2670–2675.

Loic Landrieu et Martin Simonovsky. “Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs”. In : *arXiv preprint arXiv :1711.09869* (nov. 2017).

Daniel Maturana et Sebastian Scherer. “VoxNet : A 3D convolutional neural network for real-time object recognition”. In : *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE. 2015, p. 922–928.

X. Roynard, J.-E. Deschaud et F. Goulette. “Paris-Lille-3D : a large and high-quality ground truth urban point cloud dataset for automatic segmentation and classification”. In : *ArXiv e-prints* (nov. 2017). arXiv : 1712.00032 [cs.LG].

Lyne P Tchapmi et al. “SEGCloud : Semantic Segmentation of 3D Point Clouds”. In : *arXiv preprint arXiv :1710.07563* (2017).

Hugues Thomas et al. “Semantic Classification of 3D Point Clouds with Multiscale Spherical Neighborhoods”. In : *arXiv preprint arXiv :1808.00495* (2018).