# Urban object classification with 3D Deep-Learning

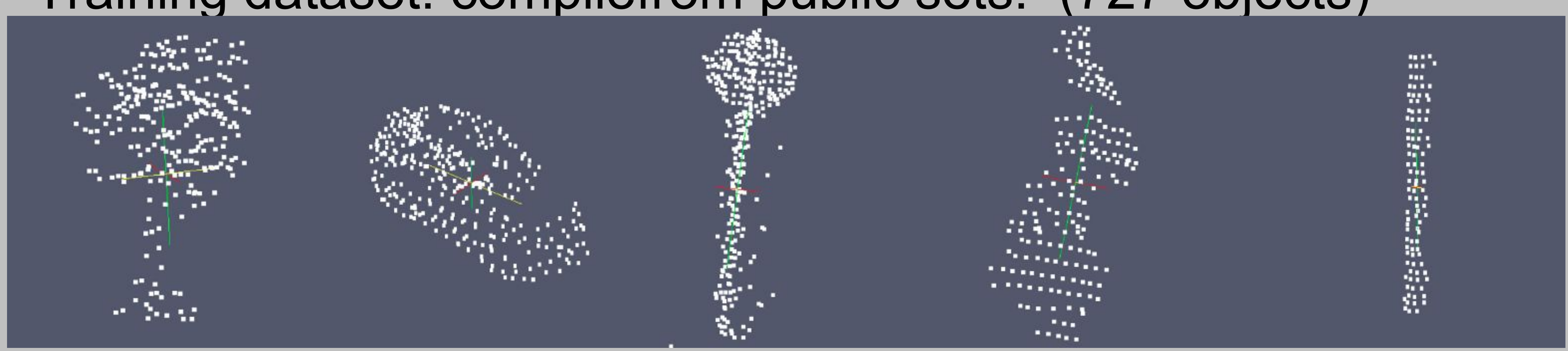
**Y. Zegaoui[1,4], M. Chaumont[1,2], G. Subsol[1], P. Borianne[3], M. Derras[4]**
[1]LIRMM, Univ. Montpellier, CNRS, Montpellier, France, [2]University of Nîmes, France,
[3]AMAP,Univ. Montpellier, CIRAD, CNRS, INRA, IRD, Montpellier, France,
[4]Berger-Levrault company, Toulouse, France
*younes.zegaoui@lirmm.fr*

## INTRODUCTION

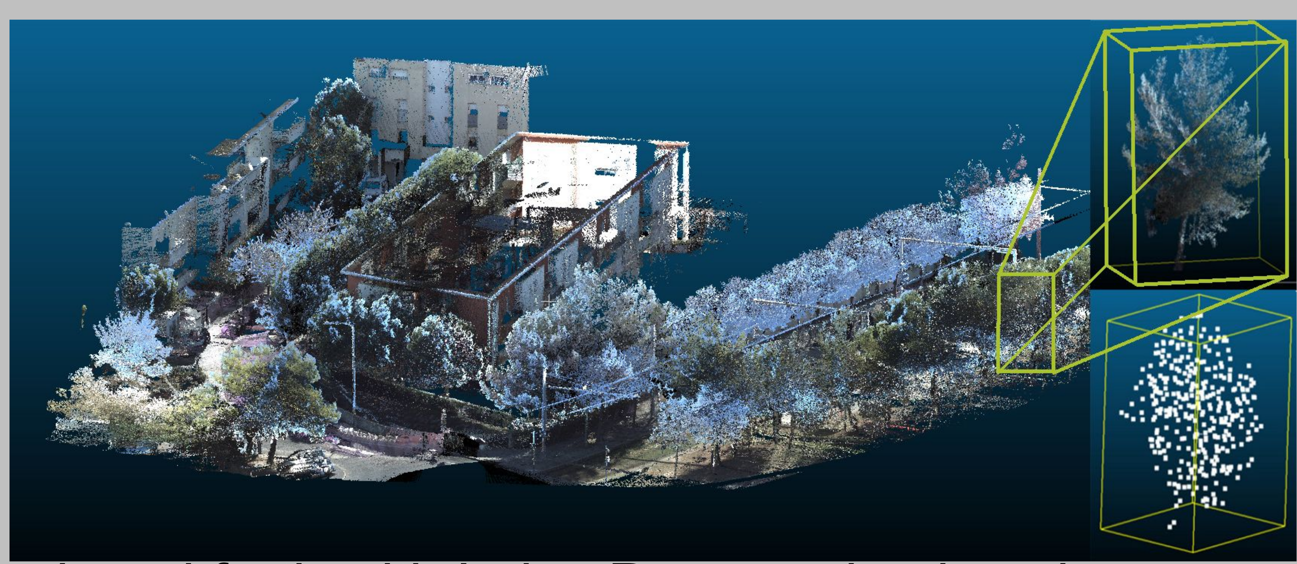
- Urban environment monitoring and management [1].
- Dynamic LiDAR acquisition to scan entire scene.
- Urban object detection and recognition.
- Deep-learning [2] techniques.
- Projecting all collected data in GIS.

## 3D URBAN OBJECT DATASETS

- 1 urban object = 1 point cloud = 512 points by uniform sampling
- 8 classes : tree, car, traffic sign, pole, pedestrian, building, noise
- Training dataset: compilefrom public sets: (727 objects)



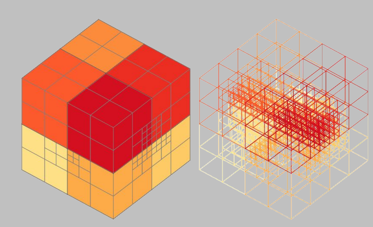
  - Sydney urban dataset:
    http://www-personal.acfr.usyd.edu.au/a.quadros/objects4.html
  - Kevin Lai dataset:
    https://sites.google.com/site/kevinlai726/datasets
  - Paris rue Madame dataset:
    http://cmm.ensmp.fr/~serna/rueMadameDataset.html
- Test dataset: our LiDAR acquisition:



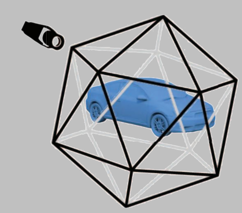
  - 200 meters back and forth with Leica Pegasus backpack
  - 27 millions points (x, y, z) + RGB
  - 174 urban objects isolated manually
    - **http://www.lirmm.fr/~zegaoui/**

## DEEP-LEARNING FOR 3D POINT CLOUD CLASSIFICATION

- Voxelizing the clouds [3]
  - Discretizing space around the cloud



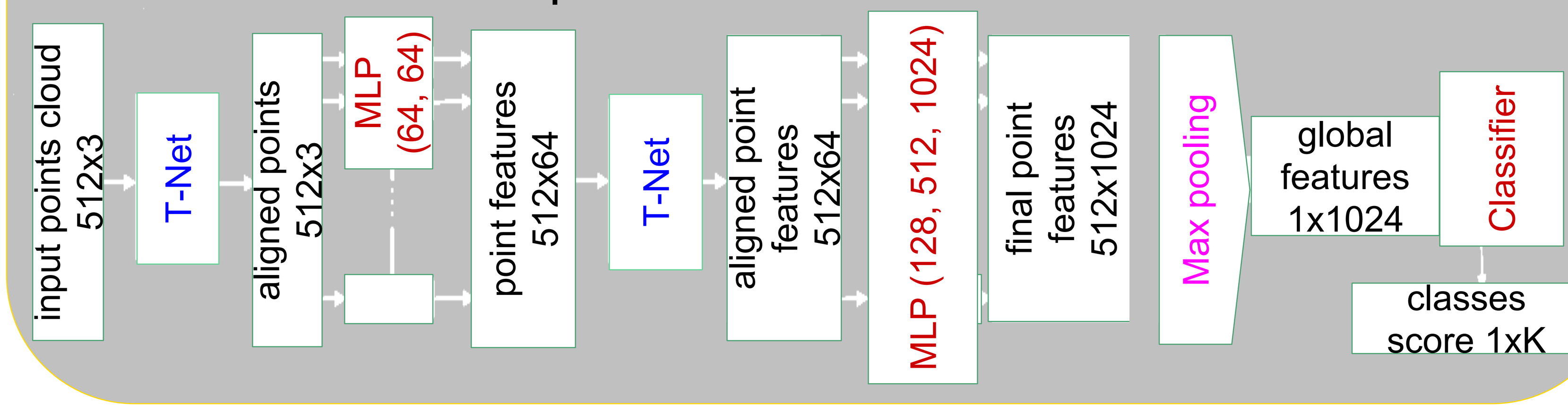
- Using multi-views [4]
  - Synthesizing 2D images from multiple point of view around the cloud



- Learning directly on point [5]

## SELECTED METHOD: POINT-NET FRAMEWORK [5]

- Points (x, y, z) are directly processed
- Coordinate frame normalized with T-Net
- Invariant w.r.t. the point order




## BASELINE CLASSIFICATION EXPERIMENT [6]

- F measure is used to evaluate the performances

| | | Ground Truth | | | | |
|---|---|---|---|---|---|---|
| | | Tree (75) | Car (39) | Traffic sign/light (8) | Pole (23) | Pedestrian (15) |
| Classification | Tree | 69 | 2 | 0 | 8 | 0 |
| | Car | 1 | 33 | 0 | 0 | 0 |
| | Traffic sign/light | 4 | 0 | 4 | 12 | 2 |
| | Pole | 0 | 0 | 3 | 1 | 0 |
| | Pedestrian | 1 | 0 | 1 | 0 | 12 |
| | Building | 0 | 0 | 0 | 2 | 0 |
| | Noise | 0 | 4 | 0 | 0 | 1 |
| F measure | | 0.896 | 0.904 | 0.267 | 0.074 | 0.828 |

## CLASSIFICATION EXPERIMENT WITH ADDITIONAL DATA

- We enrich our training set with clouds from the Paris-Lille dataset:
  http://caor-mines-paristech.fr/fr/paris-lille-3d-dataset/

| | | Ground Truth | | | | |
|---|---|---|---|---|---|---|
| | | Tree (75) | Car (39) | Traffic sign/light (8) | Pole (23) | Pedestrian (15) |
| Classification | Tree | 70 | 0 | 0 | 2 | 1 |
| | Car | 2 | 39 | 0 | 0 | 1 |
| | Traffic sign/light | 2 | 0 | 7 | 10 | 2 |
| | Pole | 0 | 0 | 1 | 11 | 0 |
| | Pedestrian | 0 | 0 | 0 | 0 | 11 |
| | Building | 1 | 0 | 0 | 0 | 0 |
| | Noise | 0 | 0 | 0 | 0 | 0 |
| F measure | | 0.946 | 0.963 | 0.467 | 0.629 | 0.815 |

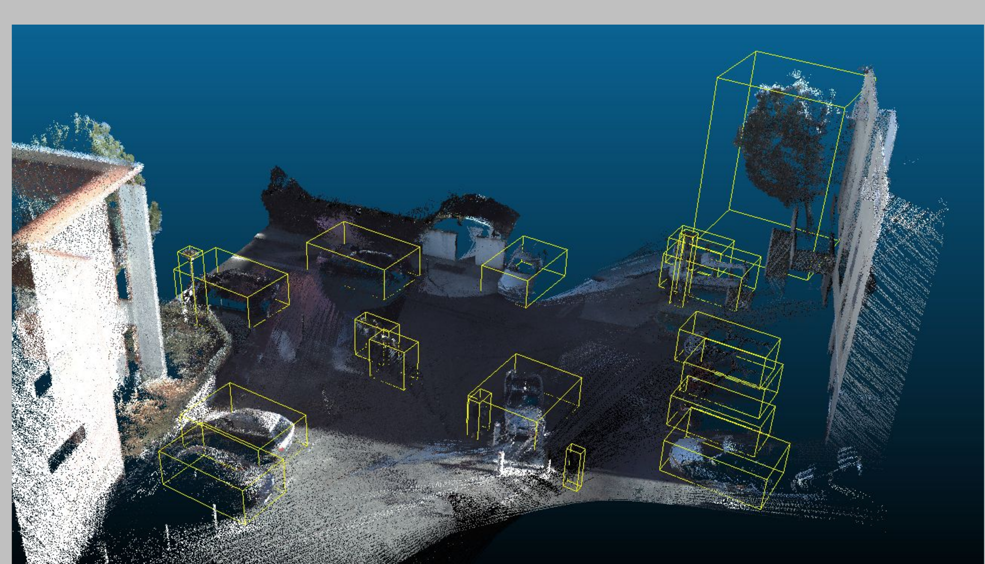
## CLASSIFICATION EXPERIMENT WITH CLASS FUSION

We fuse traffic sign/light and pole classes into TLSP class

| | | Ground Truth | | | |
|---|---|---|---|---|---|
| | | Tree (75) | Car (39) | TLSP (31) | Pedestrian (15) |
| Classification | Tree | 69 | 3 | 3 | 0 |
| | Car | 1 | 27 | 0 | 0 |
| | Pole | 2 | 0 | 28 | 5 |
| | Pedestiran | 0 | 0 | 0 | 9 |
| | Building | 1 | 0 | 0 | 0 |
| | Noise | 2 | 9 | 0 | 1 |
| F measure | | 0.920 | 0.806 | 0.849 | 0.750 |

## CONCLUSION

- Interesting results for classification despite the small size of our dataset
- Class confusion "traffic sign"/"pole" in the prediction step
- Adding training data improves results



## FUTURE WORK

- Acquisitions of new datasets
- Object localization in a scene
- 3D+t analysis of scene variation

## REFERENCES

- [1] Smart City Concept: What It Is and What It Should Be. I. Zubizarreta, A. Seravalli, S. Arrizabalaga. Journal of Urban Planning and Development, 2015
- [2] Deep learning. Y. Lecun, Y. Bengio, G. Hinton. Nature, 2015
- [3] OctNet: Learning Deep 3D Representations at High Resolutions Gernot Riegler. A. Osman, A. Geiger CVPR, 2017
- [4] Synthesizing 3D Shapes via Modeling Multi-View Depth Maps and Silhouettes with Deep Generative Networks. A. Arsalan et al. CVPR, 2017
- [5] PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. C. R. Qi, H. Su, K. Mo, L. Guibas. CVPR, 2017
- [6] First Experiments of Deep Learning on LiDAR Point Clouds for Classification of Urban Objects. Y. Zegaoui, M. Chaumont, G. Subsol, P. Borianne, M. Derras. CFPT 2018,

*We thank Leica Geosystems for its technical support as well as providing the LiDAR acquisition. This work is supported by a CIFRE grant (ANRT)*