Title: **Self-supervised representation learning for 3D point clouds**

valeo.ai

Abstract:

Over the last few years, deep learning methods on point clouds have achieved impressive results on 3D scene understanding, such as semantic segmentation. However, these methods require a large amount of training data manually annotated by humans: an expensive and time-consuming process, especially for 3D data. Yet, the size of annotated datasets for 3D scene understanding does not reach yet the scale of those available for images. A promising approach to exploit larger dataset while limiting the need for costly annotations is the use of self-supervised or unsupervised representation learning [1, 2, 3, 4], where annotation-free pretext tasks exploiting only the information available within the data permits network training

While being the object of a growing literature for image understanding, self-supervision on point clouds is only emerging [5, 6, 7, 8]. Yet, self-supervision on 3D data is of major interest for autonomous driving systems, which make an intensive use of LIDAR data.

The objective of this internship is to develop self-supervised learning methods on point clouds, exploiting 3D data specificities and side-information such as images or color often accompanying 3D data to define the pretext tasks. A possible starting direction is to adapt the recent method [2] to 3D scene datasets such as ScanNet or NuScenes while targeting a downstream task such as semantic segmentation.

Applicants : finishing their Master of Science, with a strong background in computer vision, machine (deep) learning and, ideally, 3D geometry processing.
Time: spring and summer 2021 (5-6 months)
Location: valeo.ai, 15 rue de la Baume, 75008, Paris
Main advisors: G. Puy, A. Boulc'h, S. Gidaris and A. Bursuc.
Contact: gilles.puy@valeo.com, alexandre.boulch@valeo.com

References:

[1] "Unsupervised representation learning by predicting image rotations". S. Gidaris, P. Singh, N. Komodakis, ICLR 2018.
[2] "Learning representations by predicting bags of visual words". S. Gidaris, A. Bursuc, N. Komodakis, P. Perez, M. Cord, CVPR 2020.
[3] "Momentum contrast for unsupervised representation learning". K. He, H. Fan, Yuxin Wu, Saining Xie, Ross Girshick. CVPR 2020.
[4] "A simple framework for contrastive learning of visual representations". T. Chen, S. Kornblith, Mohammad Norouzi, Geoffrey Hinton. ICML 2020.
[5] "PointContrast: Unsupervised Pre-training for 3D Point Cloud Understanding", Saining Xie, Jiatao Gu, Demi Guo, Charles R. Qi, Leonidas Guibas, Or Litany, ECCV 2020.
[6] "Self-supervised deep learning on point clouds by reconstructing space." Sauder, Jonathan, and Bjarne Sievers. Advances in Neural Information Processing Systems. 2019.
[7] "Self-supervised Learning of Point Clouds via Orientation Estimation", O. Poursaeed et al., 3DV 2020, https://arxiv.org/abs/2008.00305
[8] "OcCo: Pre-Training by Completing Point Clouds," Hanchen Wang, Qi Liu, Xiangyu Yue, Joan Lasenby, Matthew J. Kusner, https://arxiv.org/pdf/2010.01089.pdf