



**Laboratoire de Vision et
d'Apprentissage pour l'analyse de
scène**

CEA Saclay 91191 Gif-sur-Yvette France
<https://kalisteo.cea.fr>

Date de parution : 04/07/2021
Référence : LVA-2021-Th2

Contact :
Nicolas Granger nicolas.granger@cea.fr
Hervé le Borgne herve.le-borgne@cea.fr
Quoc Cuong Pham quoc-cuong.pham@cea.fr

Offre de thèse de doctorat

Méthodes faiblement-supervisées pour l'analyse de nuages de points 3D

Présentation du laboratoire d'accueil

Basé à Paris-Saclay, le CEA List est l'un des quatre instituts de recherche technologique de CEA Tech, direction de la recherche technologique du CEA. Dédié aux systèmes numériques intelligents, il contribue au développement de la compétitivité des entreprises par le développement et le transfert de technologies.

L'expertise et les compétences développées par les 800 ingénieurs-chercheurs et techniciens du CEA List permettent à l'Institut d'accompagner chaque année plus de 200 entreprises françaises et étrangères sur des projets de recherche appliquée s'appuyant sur 4 programmes et 9 plateformes technologiques. 21 start-ups ont été créées depuis 2003.

Labellisé Institut Carnot depuis 2006, le CEA List est aujourd'hui l'institut Carnot Technologies Numériques.

Le Laboratoire de Vision et Apprentissage pour l'analyse de scène (LVA) mène ses recherches dans le domaine de la Vision par Ordinateur (Computer Vision) selon quatre axes principaux :

- La reconnaissance visuelle
- L'analyse du comportement
- L'annotation intelligente et la caractérisation de données
- Les modèles de perception et de décision.

Contexte

L'utilisation de capteurs 3D LiDAR a considérablement progressé au cours des dernières années, atteignant un statut industrialisable tant en terme de capteurs que de méthodes. Dans le contexte de l'aide à la conduite ou de la conduite autonome, les données de LiDAR fournissent une information fine et fiable pour la perception de l'environnement. Ainsi, sur les tâches de détection, de suivi, ou encore de segmentation sémantique, de nombreuses méthodes démontrent des performances à l'état de l'art en utilisant cette seule source de données. Ces méthodes se caractérisent par une importante réutilisation des développées pour les images dans le domaine du spectre visible, avec trois axes principaux de contributions: la réduction du problème à la 2D, l'extension des méthodes 2D à la 3D, et l'exploitation de propriétés spécifiques à la 3D pour affiner la qualité de modélisation. Plusieurs verrous scientifiques demeurent cependant:

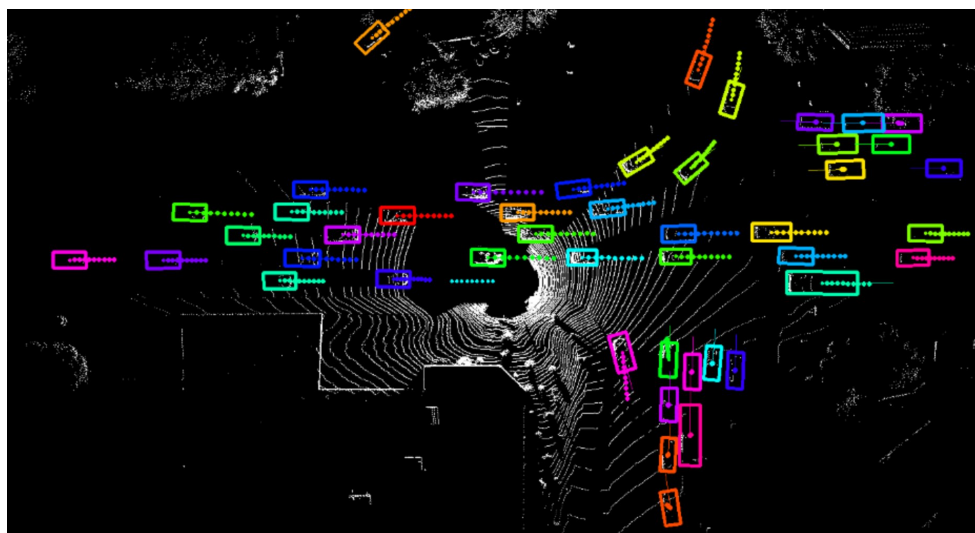
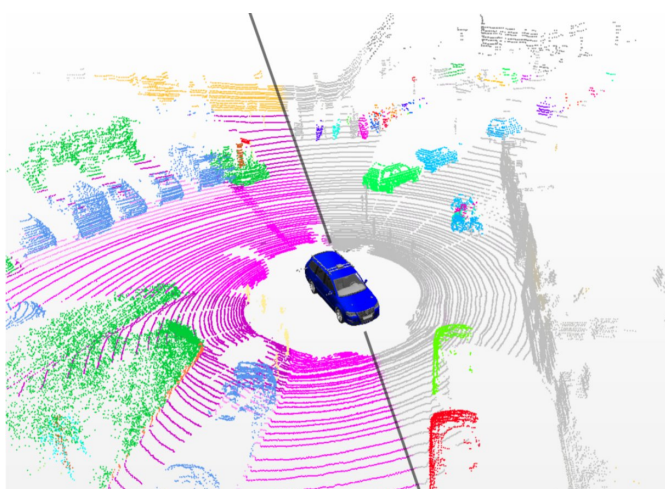
- Les bases de données publiques ne sont pas représentatives de la variété des situations et des environnements habituels de conduite, comme en témoignent les expériences de transfert de domaine [1]. C'est particulièrement le cas pour des cas rares, telles que les

situations accidentogènes. En vue d'un déploiement à grande échelle, il est donc nécessaire d'entraîner les modèles sur des données plus diverses.

- Les données d'une unique prise de vue et d'un seul capteur contiennent de multiples ambiguïtés provenant d'une part des occultations et d'autre part de la faible résolution des capteurs [5].

Parallèlement, les méthodes d'apprentissage auto-supervisées ou faiblement supervisées dans les domaines de l'image, du texte ou de la vidéo se sont considérablement améliorées ces dernières années, et adapter ces démarches à des données 3D [2] permettrait de tirer parti de larges quantités de données non-annotées qui peuvent être acquises à moindre coût.

Parmi ces approches celles s'appuyant sur la cohérence temporelle semblent particulièrement adaptées au contexte des nuages de points 3D. Les approches temporelles n'ont été que peu explorées jusqu'ici [3,4] du fait de la complexité algorithmique qu'elles impliquent, mais sont rendues plus accessibles par les nouvelles générations de calculateurs GPU. La combinaison de points issues d'une révolution du LiDAR apporterait des données plus résolues sur la scène à analyser.



Objectif de la thèse

L'objectif de cette thèse est d'améliorer la fiabilité et la qualité de généralisation des modèles de perception sur des nuages de points 3D. La personne recrutée pourra dans un premier temps étudier la pertinence des méthodes développées sur d'autres types de données puis chercher à tirer parti des hypothèses spécifiques aux données 3D, notamment la cohérence temporelle de l'environnement observé.

Références

- [1] Saltori, C., Lathuilière, S., Sebe, N., Ricci, E. and Galasso, F., (2020), SF-UDA 3D: Source-Free Unsupervised Domain Adaptation for LiDAR-Based 3D Object Detection. In 2020 International Conference on 3D Vision (3DV) (pp. 771-780).
- [2] Xie, S., Gu, J., Guo, D., Qi, C.R., Guibas, L. and Litany, O., (2020), August. Pointcontrast: Unsupervised pre-training for 3d point cloud understanding. In European Conference on Computer Vision (pp. 574-591). Springer, Cham.
- [3] Choy, C., Gwak, J. and Savarese, S., (2019). 4d spatio-temporal convnets: Minkowski convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3075-3084).
- [4] Luo, W., Yang, B. and Urtasun, R., (2018). Fast and furious: Real time end-to-end 3d detection, tracking and motion forecasting with a single convolutional net. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 3569-3577).
- [5] Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G. and Beijbom, O., (2020). nuscenes: A multimodal dataset for autonomous driving. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11621-11631).

Profil du(de la) candidat(e) :

- Master recherche ou ingénieur avec une expérience en vision et en machine learning
- Compétences : vision par ordinateur, apprentissage automatique, notamment *deep learning*
- Programmation : Python, C++, frameworks de Deep Learning (Tensorflow, PyTorch)

Si vous vous reconnaissez dans ces compétences et ce profil et si développer vos compétences sur la thématique de l'IA, au sein d'un institut ambitieux et reconnu, au cœur de l'environnement dynamique du plateau de Saclay s'inscrit dans votre projet professionnel, merci de transmettre CV + lettre de motivation à :

nicolas.granger@cea.fr / herve.le-borgne@cea.fr / quoc-cuong.pham@cea.fr