



CEA List
Service d'Intelligence Artificielle pour le Langage et la Vision
Centre de Saclay 91191 Gif-sur-Yvette France
<http://www.kalisteo.eu>

Contact Florian Chabot
Tél +33 (0)1 69 08 02 88
E-mail florian.chabot@cea.fr

STAGE 2021

Réf : LVA-21-S4

Prédiction de cartes de profondeur à partir d'images monoculaires

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 :

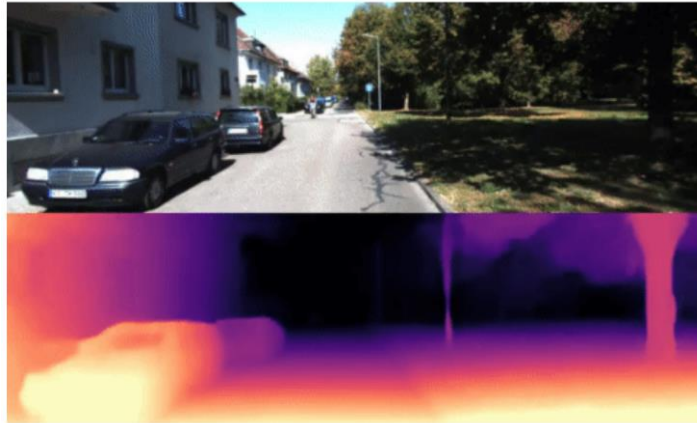
- La reconnaissance visuelle (détection et/ou segmentation d'objets, de personnes, de patterns ; détection d'anomalies ; caractérisation)
- L'analyse du comportement (reconnaissance de gestes, d'actions, d'activités, de comportements anormaux ou spécifiques pour des individus, un groupe, une foule)
- Annotation intelligente (annotation à grande échelle de données visuelles 2D/3D de manière semi-automatique)
- Perception et décision (processus de décision markovien, navigation).

Description du stage

L'objectif de ce stage est de proposer une méthode hybride (basée sur l'apprentissage profond supervisé et auto-supervisé) pour la prédiction de cartes de profondeur à partir d'une image pour des applications de type véhicules autonomes. Les approches supervisées récentes pour résoudre cette tâche utilisent généralement un nuage de points 3D (LIDAR) comme vérité terrain pour superviser l'apprentissage [1, 2]. Cependant, l'information provenant du LIDAR est incomplète, sparse et parfois bruitée ce qui ne permet pas de prédire avec précision la profondeur sur certaines zones de l'image. Une autre famille de méthodes s'intéresse à l'analyse du mouvement entre deux images pour inférer la profondeur par des mécanismes d'apprentissage auto-supervisé [3]. Ces approches ont l'avantage de n'utiliser que des données non-annotées pour entraîner le modèle. Cependant, elles présentent deux défauts majeurs :

- l'incapacité d'apprendre correctement la profondeur quand les objets de la scène sont en mouvement
- la profondeur est estimée à un facteur d'échelle prêt

Pour pallier ces problèmes, le stagiaire devra dans un premier temps investiguer l'ajout de l'auto-supervision dans les approches supervisées afin de d'améliorer l'estimation de la profondeur. Dans un second temps, il devra analyser les cas d'erreurs et proposer des solutions pour les résoudre. Enfin, l'algorithme développé devra être comparé aux méthodes existantes sur des benchmarks académiques.



Estimation de profondeur monoculaire par apprentissage auto-supervisé [3]

Références

- [1] Fu et al. Deep Ordinal Regression Network for Monocular Depth Estimation, CVPR 2018
- [2] Diaz et al. Soft Labels for Ordinal Regression, CVPR 2019
- [3] Godard et al. Digging into Self-Supervised Monocular Depth Prediction, ICCV 2019

Niveau demandé :	Ingénieur, Master 2
Ce stage ouvre la possibilité de poursuite en thèse et ingénieur R&D dans notre laboratoire.	
Durée :	6 mois
Rémunération :	entre 700 € et 1300 € suivant la formation.
Compétences requises :	
<ul style="list-style-type: none"> - Vision par ordinateur - Apprentissage automatique (deep learning) - Reconnaissance de formes - C/C++, Python - La maîtrise d'un framework d'apprentissage profond (en particulier Tensorflow ou PyTorch) est un plus. 	