CEA List
Laboratoire de Vision et d'Ingénierie des Contenus
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 2019 Réf : AS-19-S4

Segmentation d'instances par apprentissage profond

Présentation du laboratoire d'accueil

Au cœur du Plateau de Saclay (Ile-de-France), l'institut CEA LIST focalise ses recherches sur les systèmes numériques intelligents. Porteurs d'enjeux économiques et sociétaux majeurs, ses programmes de R&D sont centrés sur les systèmes interactifs (intelligence ambiante), les systèmes embarqués (architectures, ingénierie logicielle et systèmes), les capteurs et le traitement du signal (contrôle industriel, santé, sécurité, métrologie).

Dédiés à la recherche technologique, les 700 ingénieurs-chercheurs et techniciens de l'institut ont pour objectif de favoriser l'innovation et son transfert autour de partenariats industriels pérennes. La culture projet et l'excellence scientifique des équipes de l'Institut sont au cœur de cette ambition.

Le Laboratoire de Vision et d'Ingénierie des Contenus mène ses recherches dans les domaines de la Vision par Ordinateur (Computer Vision) et l'analyse automatique de texte avec le défi d'extraire et d'organiser l'information à partir de documents faiblement ou non structurés (texte, image, vidéo, réseaux de capteurs).

Contexte du stage

Ce stage s'inscrit dans les activités d'Analyse de Scène du Laboratoire Vision et Ingénierie des Contenus du CEA List. Le laboratoire développe dans le cadre de projets R&D, des systèmes d'analyse automatique d'images et de vidéo intégrant des technologies avancées de détection, localisation, classification et suivi temporel d'objets d'intérêt, de reconnaissance d'activité et de comportements, ainsi que des outils de modélisation automatique de la scène.

Description du stage

L'application principale visée par ce stage est l'analyse de scènes routières, notamment pour des applications de conduite autonome. La segmentation d'instances est un problème complexe qui se situe entre la détection d'objets (prédiction des boîtes englobantes des instances de la scène) et la segmentation sémantique (attribution d'une classe à chaque pixel de l'image). Un algorithme de segmentation d'instances est donc capable de produire un masque pour chaque objet d'intérêt de l'image (cf Figure).

Exemple des résultats obtenus avec Mask R-CNN [1] sur le dataset Cityscapes [4]

CEA List
Laboratoire de Vision et d'Ingénierie des Contenus
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

Depuis l'avènement des réseaux de neurones profonds convolutifs (CNN), un certain nombre de méthodes ont été proposées afin de solutionner cette problématique, notamment l'approche par détection Mask-RCNN [1]. Ce système très performant n'est néanmoins pas compatible avec les contraintes temps réel que requièrent les pipelines de conduite autonome.

Dans le cadre de ce stage de recherche, le stagiaire aura pour première mission d'analyser les avantages et les inconvénients des approches récentes de la littérature [1, 2, 3]. Dans un second temps, il devra proposer une méthode permettant de pallier une ou plusieurs faiblesses des approches de l'état de l'art. Cette méthode devra être validée par une évaluation quantitative sur plusieurs bases de données publiques consacrées à la segmentation d'instances (cityscape [4], kitti [5]). Enfin, il sera amené à intégrer sa solution dans les différents algorithmes du laboratoire.

Références

[1] "Mask R-CNN", Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick, ICCV 2017

[2] "Box2Pix: Single-Shot Instance Segmentation by Assigning Pixels to Object Boxes", IEEE Intelligent Vehicles Symposium 2018

[3] "Path Aggregation Network for Instance Segmentation", Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, Jiaya Jia, CVPR 2018

[4] "The Cityscapes Dataset for Semantic Urban Scene Understanding", M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, CVPR 2016

[5] "Object Scene Flow for Autonomous Vehicles", Moritz Menze and Andreas Geiger, CVPR 2015

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 : - Vision par ordinateur - Une bonne connaissance des algorithmes d'apprentissage et des réseaux de neurones (Deep Learning) - Compétences informatique : C++, python, linux/gcc, cmake, git - Une première expérience d'utilisation d'outils de deep learning (Pytorch, Caffe, Tensorflow) est un plus	