



STAGE 2020

Réf : LVA-2020-S8

Segmentation interactive d'objets par réseaux de neurones profonds

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 et de l'Intelligence Artificielle selon quatre axes principaux :

- La reconnaissance visuelle (détection, reconnaissance fine et segmentation d'objets, segmentation de scène, détection d'anomalies)
- L'analyse du comportement (reconnaissance de gestes, d'actions, d'activités, détection de comportements anormaux ou spécifiques)
- L'annotation intelligente à grande échelle de données visuelles
- Les modèles de perception et décision.

Description du stage

La détection fine d'objets dans une image ou vidéo est une problématique très étudiée dans le domaine de la vision par ordinateur. Cette technique est utilisée dans de nombreux domaines d'application, tels que l'imagerie médicale, la vidéosurveillance et la conduite autonome.

La plupart des méthodes de détection sont fondées sur l'apprentissage supervisé et semi-supervisé. Elles requièrent cependant une quantité importante de données annotées, en particulier pour les méthodes basées sur les réseaux de neurones profonds (Deep Learning). Or, l'annotation manuelle des objets d'intérêt dans une image est une tâche longue et coûteuse, notamment pour l'annotation fine (au niveau pixelique ou polygonal).

L'objectif principal de ce stage est de concevoir des algorithmes d'annotation semi-automatique et précise d'objets nécessitant très peu d'effort humain (e.g. simple clic).

Une première approche consistera à segmenter chaque objet pré-localisé par un opérateur (e.g. un pixel de l'objet, un tracé simple, une boîte englobante, ...). Cette segmentation d'instance consiste à donner un label unique pour tous les pixels de chaque objet. Le stagiaire aura pour première mission d'étudier les approches de l'état de l'art, et de concevoir une nouvelle méthode semi-automatique de segmentation optimisant l'apport de l'opérateur. Il effectuera une étude comparative avec les méthodes existantes [1, 2, 3] (voir figure).

Une deuxième phase du stage consistera à étendre l'algorithme conçu à la correction interactive de correction de segmentation [4].

Enfin, les algorithmes développés seront intégrés à la chaîne d'outils d'annotation du laboratoire.



Exemple d'annotation fine polygonale sur le dataset Cityscapes, générées semi-automatiquement (gauche) et manuellement (droite) [4]

Mots-clés :

vision par ordinateur, apprentissage profond, segmentation interactive, annotation.

Références

- [1] He, Gkioxari, Dollár, Girshick. Mask R-CNN. ICCV, 2017.
- [2] Rajchl, Lee, Oktay, Kamnitsas, Passerat-Palmbach, Bai, Damodaram, Rutherford, Hajnal, Kainz, Rueckert. DeepCut: Object Segmentation from Bounding Box Annotations using Convolutional Neural Networks. IEEE, 2016.
- [3] Acuna, Ling, Kar, Fidler. Efficient Interactive Annotation of Segmentation Datasets with Polygon-RNN++. CVPR, 2018.
- [4] Benenson, Popov, Ferrari. Large-scale interactive object segmentation with human annotators. CVPR, 2019.

Profil du candidat/de la candidate

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 - 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.	