

Suivi visuel par segmentation d'instances par apprentissage profond

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)
- L'annotation intelligente (annotation à grande échelle de données visuelles 2D/3D de manière semi-automatique)
- La perception et la décision (processus de décision markovien, navigation).

Description du stage



Figure 1. Exemples de résultats de suivi de segmentation d'instance avec la méthode CINM [1] sur les datasets DAVIS 2016 [2], Youtube-Objects [3] et SegTrack v2 [4].

Le suivi d'objets dans une image 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 la vidéosurveillance et la conduite autonome. Nous pouvons distinguer deux différentes familles d'algorithmes de suivi d'objets: le suivi « par détection » qui repose sur une association a posteriori d'instances détectées image par image et le suivi « visuel » qui propage directement la détection d'un objet d'une image à l'autre. Le laboratoire travaillant sur l'analyse de vidéos de sports collectifs, nous avons besoin de suivre des objets particulièrement difficiles à détecter image par image tels des balles ou des joueurs.



Figure 2. Illustration du suivi de balles de volleyball¹

L'objectif de ce stage est donc de développer un algorithme de suivi visuel d'instances robuste aux occultations partielles voire intégrales souvent retrouvées dans les scènes de sport collectif. Une première approche pourra reposer sur un suivi visuel de boîtes englobantes autour d'un objet d'intérêt cible. Le stagiaire aura pour première mission d'étudier les approches de l'état de l'art, et de concevoir une nouvelle méthode dans le cadre de vidéos de sport collectif. Une deuxième phase du stage consistera à étendre l'algorithme conçu au suivi visuel pixellique de la cible (« segmentation d'instance »). Il pourra aussi effectuer une étude comparative avec les méthodes existantes sur des bases publiques [1].

[1] L. Bao, B. Wu and W. Liu. CNN in MRF: Video Object Segmentation via Inference in A CNN-Based Higher-Order Spatio-Temporal MRF (2018)

[2] F. Perazzi, J. Pont-Tuset, B. McWilliams, L. Van Gool, M. Gross, and A. Sorkine-Hornung. A benchmark dataset and evaluation methodology for video object segmentation. In CVPR (2016)

[3] S. D. Jain and K. Grauman. Supervoxel-consistent foreground propagation in video. In ECCV, 2014

[4] F. Li, T. Kim, A. Humayun, D. Tsai, and J. M. Rehg. Video segmentation by tracking many figure-ground segments. In ICCV (2013)

¹ <https://towardsdatascience.com/ball-tracking-in-volleyball-with-opencv-and-tensorflow-3d6e857bd2e7>

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. 	