

STAGE 2020

Réf : LVA-2020-S2

Reconnaissance d'interactions dans des vues egocentriques

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

L'objectif de ce stage est d'aborder l'analyse des flux vidéo dit égocentriques, c'est-à-dire filmés à la première personne. Ce type de point de vue est particulièrement utile dans le domaine industriel pour l'aide au contrôle qualité, celui de la robotique ou de l'aide aux personnes déficientes visuelles. Les principales problématiques sont l'analyse des interactions de la personne avec les objets et l'analyse fine des actions réalisées. La diversité et les complexités des objets manipulés, des gestes, des contextes sont autant de difficultés à appréhender.

Pour cela, le stage s'appuiera d'une part sur l'état de l'art actuel sur la reconnaissance d'action [1] et d'autre part sur les technologies du laboratoire d'analyse de l'activité [2], d'analyse d'interaction [3] ainsi que d'analyse de posture humaine [4].

Afin de s'intéresser à un exemple concret, le stage se focalisera sur le dataset EPIC KITCHENS [5] qui est constitué de vidéos d'actions courantes de cuisine et d'entretien réalisées dans différentes cuisines. Une solution de reconnaissance d'action basée deep learning sera mise en place et évaluée selon le benchmark du dataset. La mise en place d'un démonstrateur temps réel pourra aussi être envisagé selon l'avancée du stage.



Exemple d'images issues du dataset EPIC KITCHEN



Mots-clés :

vision par ordinateur, apprentissage profond, reconnaissance d'interactions, vues egocentriques.

Références

- [1] Wang, X., Wu, Y., Zhu, L., Yang, Y. Baidu-UTS Submission to the EPIC-Kitchens Action Recognition Challenge 2019, ArXiv, 2019.
- [2] Vaquette, G., Orcesi, A., Lucat, L., & Achard, C. (2017, May). The DAily Home Life Activity Dataset: A High Semantic Activity Dataset for Online Recognition. In Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference on (pp. 497-504).
- [3] Chafik, S., Orcesi, A., Audigier, R., Luvison, B. , Classifying All Interacting Pairs in a Single Shot, ArXiv 2019 (à paraître)
- [4] Benzine, A., Luvison, B., Pham, Q.C., Achard, C., Deep, robust, and single-shot 3D multi-person human pose estimation from monocular images, ICIP 2019.
- [5] EPIC Kitchens dataset, <https://epic-kitchens.github.io/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 : <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.	