

Estimation de postures 3D avec des contraintes temporelles 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 et de l'Intelligence Artificielle selon quatre axes principaux :

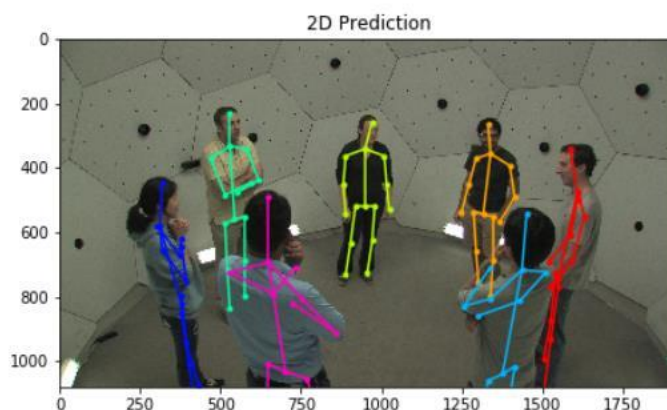
- La reconnaissance visuelle (détection, reconnaissance fine et segmentation d'objets, segmentation de scène, détection d'anomalies)
- L'analyse du comportement (reconnaissance de gestes, d'actions, d'activités, détection de comportements anormaux ou spécifiques)
- L'annotation intelligente à grande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

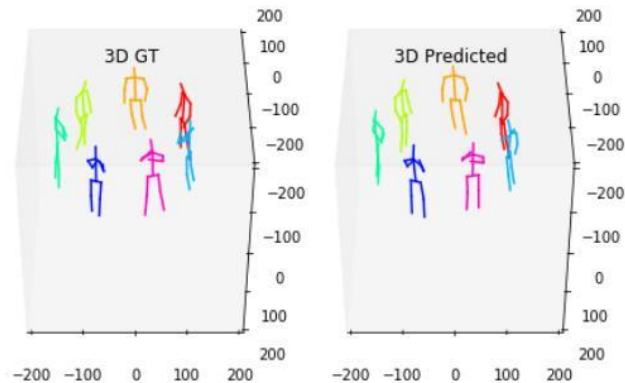
Description du stage

Comprendre automatiquement l'activité humaine se déroulant dans un flux vidéo est l'un des principaux challenges actuels de la vision par ordinateur. Bien que le sujet soit abordé depuis longtemps, les applications qui en résultent ne se limitent aujourd'hui qu'à des contextes et des gestes bien particuliers. L'un des prérequis à cette analyse est généralement l'extraction précise de la posture 3D des personnes présentes dans la scène, afin d'avoir une représentation beaucoup plus compacte de l'état des personnes à un instant t.

Grâce aux méthodes basées sur les réseaux de neurones profonds (deep learning), des progrès notables ont été réalisés sur cette tâche. Néanmoins la plupart des approches actuelles sont limitées à l'estimation de la posture 3D d'une seule personne et n'utilisent pas d'information temporelle afin de rendre plus cohérente et plus robuste l'estimation de la posture.

L'objectif du stage sera d'étendre un algorithme basé deep learning développé au sein du laboratoire [1] afin de permettre l'estimation de postures 3D de plusieurs personnes à partir de séquences d'images.





Exemple d'estimation de posture 3D multi-personnes réalisé sur une image du dataset Panoptic

Mots-clés :

vision par ordinateur, apprentissage profond, estimation de posture 3D, multi-personnes, modèle temporel.

Références

[1] 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.

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. 	