

STAGE 2022

Réf : LVA-22-S1

IA frugale pour la ré-identification d'objets par l'adaptation de domaine non supervisée

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.

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 :

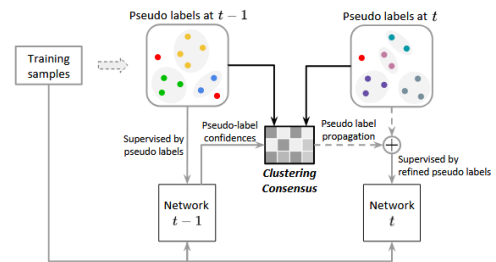
- La reconnaissance visuelle (détection et/ou segmentation d'objets, de personnes, de patterns ; détection d'anomalies ; caractérisation)
- L'analyse du comportement (reconnaissance de gestes, d'actions, d'activités, de comportements anormaux ou spécifiques pour des individus, un groupe, une foule)
- L'annotation intelligente (annotation à grande échelle de données visuelles 2D/3D de manière semi-automatique)
- Les modèles de perception pour l'aide à la décision.

Description du stage

La ré-identification automatique d'objets (personnes ou voitures par exemple) vus par des caméras est une fonctionnalité clé pour les applications de vidéo protection ou encore de surveillance routière. Ce problème en ensemble ouvert consiste à retrouver les occurrences d'un objet dans un ensemble d'images, problème que l'on retrouve également dans des applications de recherche d'images ou d'instances. Malgré les nombreux travaux sur la problématique ces dernières années, la modélisation de l'apparence d'un objet reste un défi. En effet, elle doit pouvoir discriminer des objets distincts (malgré leurs éventuelles similitudes) tout en étant robuste face à la forte variabilité de leur apparence visuelle (due aux postures si on s'intéresse à des personnes, aux points de vue, aux conditions d'illumination, à la sensibilité de la caméra, à sa résolution, ...). Si les méthodes d'apprentissage profond supervisé ont fortement amélioré les performances de ré-identification sur certains jeux de données académiques, leur mise en œuvre dans un contexte opérationnel demeure difficile. En effet, un modèle appris sur un jeu de données est très souvent peu performant s'il est appliqué tel quel sur d'autres jeux de données. Par ailleurs, annoter manuellement les données du domaine cible est une tâche fastidieuse donc coûteuse.

Nous nous intéresserons dans ce stage au problème d'adaptabilité des modèles d'apparence visuelle à un domaine cible dont on ne possède que des données non annotées. Cette adaptation de domaine non supervisée vise à transférer les connaissances acquises dans un domaine source annoté afin de mesurer correctement les affinités entre les instances dans le domaine cible [1,2,3,4]. L'influence de l'ensemble de données source peut varier selon les méthodes proposées, allant même jusqu'à être écartée dans les dernières approches complètement non supervisées [5,6,7].

L'objectif du stage est, dans un premier temps, d'étudier plusieurs méthodes de l'état de l'art. Le candidat devra les évaluer afin d'apprécier leurs avantages et leurs limitations. Dans un deuxième temps, le candidat devra étudier dans quelles mesures les méthodes auto-supervisées viennent supplanter les approches par adaptation de domaine non supervisée. Le candidat sera également invité à proposer des améliorations aux méthodes de l'état de l'art pour pallier un ou plusieurs problèmes identifiés. Les travaux menés durant le stage pourront faire l'objet de publications scientifiques.



Exemples d'une méthode de réidentification d'objets par adaptation de domaine non supervisée à gauche [1] et d'une méthode de réidentification d'objets complètement non supervisée à droite [5].

Keywords

Object re-identification, image retrieval, deep learning, unsupervised domain adaptation, self-supervised learning, frugal AI.

Références

- [1] Isobe, T., Li, D., Tian, L., Chen, W., Shan, Y., & Wang, S. (2021). Towards discriminative representation learning for unsupervised person re-identification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 8526-8536).
- [2] Dubourvieux, F., Audigier, R., Loesch, A., Ainouz, S., & Canu, S. (2021, January). Unsupervised domain adaptation for person re-identification through source-guided pseudo-labeling. In *2020 25th International Conference on Pattern Recognition (ICPR)* (pp. 4957-4964). IEEE.
- [3] Ge, Y., Chen, D., & Li, H. (2019, September). Mutual Mean-Teaching: Pseudo Label Refinery for Unsupervised Domain Adaptation on Person Re-identification. In *International Conference on Learning Representations*.
- [4] Song, L., Wang, C., Zhang, L., Du, B., Zhang, Q., Huang, C., & Wang, X. (2020). Unsupervised domain adaptive re-identification: Theory and practice. *Pattern Recognition*, 102, 107173.
- [5] Zhang, X., Ge, Y., Qiao, Y., & Li, H. (2021). Refining Pseudo Labels with Clustering Consensus over Generations for Unsupervised Object Re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3436-3445).
- [6] Ge, Y., Zhu, F., Chen, D., Zhao, R., & Li, H. (2020). Self-paced contrastive learning with hybrid memory for domain adaptive object re-id. In *Adv. Neural Inform. Process. Syst.*, 2020
- [7] Chen, H., Wang, Y., Lagadec, B., Dantcheva, A., & Bremond, F. (2021). Joint Generative and Contrastive Learning for Unsupervised Person Re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2004-2013).

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 - Python, C/C++ - Maîtrise d'un framework d'apprentissage profond (en particulier Tensorflow ou PyTorch) 	