

Sujet de stage (Master 2) :

Self-supervised Learning for Few-shot Object Detection

Durée : 5 mois.

Lieu : Laboratoire de Traitement et Transport de l'Information (L2TI, UR 3043)
Université Sorbonne Paris Nord (USPN)

Compétences requises :

Formation initiale en traitement d'image et apprentissage.

Bonnes connaissances en programmation Python et en particulier du framework PyTorch.

Bonne capacité d'analyse et aisance en lecture d'articles scientifiques.

Mots clés : Détection d'objets, Few-shot learning, Self-supervised learning, Deep learning.

Encadrants :

A. MOKRAOUI (L2TI, anissa.mokraoui@univ-paris13.fr)

P. LE JEUNE (L2TI, pierre.lejeune@univ-paris13.fr).

Description succincte du sujet

La détection d'objets dans des images est une tâche difficile mais réalisable avec des techniques basées sur le deep learning [1, 2]. À condition de disposer d'une base de données suffisamment grande et annotée, des performances très satisfaisantes peuvent être atteintes. Lorsque cela n'est pas le cas, il devient alors très difficile d'automatiser la détection. Des techniques de few-shot learning proposent des solutions pour apprendre à partir de peu de données [3]. Ces méthodes nécessitent tout de même des bases de données annotées pour le pré-entraînement du réseau et notamment pour la partie extraction de caractéristiques. Récemment, le self-supervised learning a proposé des méthodes d'apprentissage ne nécessitant pas d'annotation permettant l'apprentissage de caractéristiques de très bonnes qualités, qui semblent bénéfiques pour de multiples tâches [4, 5]. Ces méthodes sont souvent qualifiées de contrastive car elles tentent d'agglomérer les images visuellement similaires dans un espace de représentation tandis que les images distinctes sont éloignées les unes des autres. L'application de ce genre de techniques pour la détection few-shot est prometteuse, cela apporterait un entraînement plus rapide et de meilleure qualité. Il s'agira au cours de ce stage d'utiliser et d'évaluer ce genre de techniques dans le cadre de la détection few-shot.

References

- [1] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28:91–99, 2015.
- [2] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.
- [3] Bingyi Kang, Zhuang Liu, Xin Wang, Fisher Yu, Jiashi Feng, and Trevor Darrell. Few-shot object detection via feature reweighting. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8420–8429, 2019.
- [4] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020.
- [5] Mathilde Caron, Ishan Misra, Julien Mairal, Priya Goyal, Piotr Bojanowski, and Armand Joulin. Unsupervised learning of visual features by contrasting cluster assignments. *arXiv preprint arXiv:2006.09882*, 2020.