CEA List
Laboratoire de Vision et d'Ingénierie des Contenus
Centre de Saclay 91191 Gif-sur-Yvette France
http://www.kalisteo.eu

Contact Romaric AUDIGIER
Tél +33 (0)1 69 08 01 06
E-mail romaric.audigier@cea.fr

STAGE 2019 Réf : AS-19-S8

Few-shot learning : Apprendre à apprendre à partir de peu d'exemples

Présentation du laboratoire d'accueil

Au cœur du Plateau de Saclay (Ile-de-France), l'institut CEA LIST focalise ses recherches sur les systèmes numériques intelligents. Porteurs d'enjeux économiques et sociétaux majeurs, ses programmes de R&D sont centrés sur les systèmes interactifs (intelligence ambiante), les systèmes embarqués (architectures, ingénierie logicielle et systèmes), les capteurs et le traitement du signal (contrôle industriel, santé, sécurité, métrologie).

Dédiés à la recherche technologique, les 700 ingénieurs-chercheurs et techniciens de l'institut ont pour objectif de favoriser l'innovation et son transfert autour de partenariats industriels pérennes. La culture projet et l'excellence scientifique des équipes de l'Institut sont au cœur de cette ambition.

Le Laboratoire de Vision et d'Ingénierie des Contenus mène ses recherches dans les domaines de la Vision par Ordinateur (Computer Vision) et l'analyse automatique de texte avec le défi d'extraire et d'organiser l'information à partir de documents faiblement ou non structurés (texte, image, vidéo, réseaux de capteurs).

Contexte du stage

Ce stage s'inscrit dans les activités d'Analyse de Scène du Laboratoire Vision et Ingénierie des Contenus du CEA List. Le laboratoire développe dans le cadre de projets R&D, des systèmes d'analyse automatique d'images et de vidéo intégrant des technologies avancées de détection, localisation, classification et suivi temporel d'objets d'intérêt, de reconnaissance d'activité et de comportements, ainsi que des outils de modélisation automatique de la scène.

Description du stage

De nombreuses tâches en vision par ordinateur sont aujourd'hui réalisées avec succès par les méthodes d'apprentissage profond (*deep learning*). C'est le cas par exemple de la détection et la reconnaissance d'objets dans les images, la classification d'images, la reconnaissance de personnes ou d'activités… utiles dans divers domaines d'application (vidéosurveillance, conduite autonome, imagerie médicale, etc.). L'inconvénient de ces approches basées réseaux de neurones profonds est que leur entrainement supervisé requiert de grandes quantités de données annotées. D'une part, l'annotation manuelle de données est une tâche longue et couteuse. D'autre part, ces données peuvent être rares ou difficiles à collecter. Il est alors nécessaire d'avoir des méthodes d'apprentissage qui se basent sur très peu d'exemples. Les approches de *few-shot learning* qui font partie du paradigme général de *meta-learning* (« apprendre à apprendre ») répondent à ce besoin (cf. Fig.1). Leur but est d'entrainer un modèle sur une variété de tâches d'apprentissage de sorte qu'il puisse ensuite apprendre de nouvelles tâches rapidement, précisément et en utilisant très peu d'exemples d'entrainement. Le défi est alors d'obtenir une grande capacité de généralisation du modèle face à ces nouvelles tâches.

L'objectif de ce stage est d'explorer plusieurs paradigmes de méta-apprentissage qui permettent d'apprendre à partir de peu d'exemples. En particulier, le but sera de concevoir des algorithmes de reconnaissance d'objets capables d'apprendre rapidement de nouvelles classes d'objets dont on n'a que très peu d'exemples. Plusieurs approches de l'état de l'art en *meta-learning* seront étudiées, notamment celles liées au paradigme de mémoire augmentée (*Memory Augmented Neural Networks* [1]), celles basées sur l'apprentissage de métriques [2,3,4,5] (cf. Fig.2) ou encore celles qui reposent sur l'apprentissage d'un optimiseur (*meta-learner*) [6,7,8,9] (cf. Fig.3). Dans le cadre de la reconnaissance d'objets, une nouvelle méthode sera proposée dans le but d'optimiser la capacité du modèle à effectuer de nouvelles tâches rapidement (reconnaître de nouvelles classes). Une évaluation comparative des méthodes développées sera effectuée sur divers jeux de données afin de quantifier l'amélioration des performances, la dépendance au nombre d'échantillons, ainsi que la capacité de généralisation face aux types de données. Pour finir, une extension de la méthode vers de la détection fine d'objets pourra être étudiée.

CEA List
Laboratoire de Vision et d'Ingénierie des Contenus
Centre de Saclay 91191 Gif-sur-Yvette France
http://www.kalisteo.eu

Contact Romaric AUDIGIER
Tél +33 (0)1 69 08 01 06
E-mail romaric.audigier@cea.fr

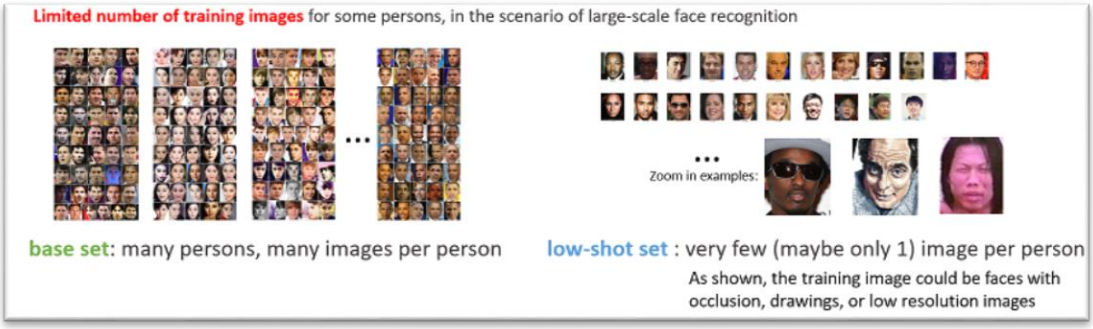

Fig. 1 Le *Low-shot Learning Challenge* [11] a pour but de reconnaitre des visages de personnes en les ayant vus sur un nombre très limité d'images.

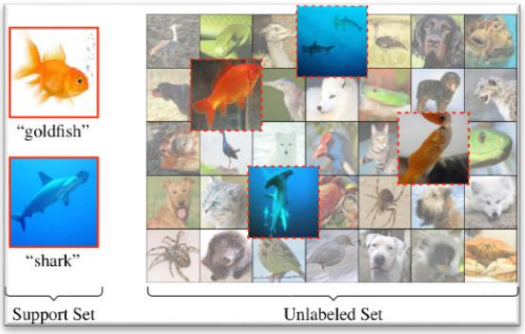

Fig. 2 La méthode [5] propose une approche semi-supervisée : Apprendre à reconnaitre de nouvelles classes d'objets à partir de peu d'exemples annotés et des données non annotées, dont de multiples distracteurs.

Fig. 3 *Dataset* Omniglot [10]. La méthode MAML [6] de *meta-learning* par apprentissage d'un optimiseur, après avoir été entraînée sur certains alphabets, est capable de reconnaitre avec une précision de 95.8% 20 nouveaux caractères d'alphabet dont un seul exemplaire est donné.

Références

[1] Santoro A. et al. (2016) One-shot Learning with Memory-Augmented Neural Networks. *arXiv.*

[2] Koch G. et al. (2015), Siamese Neural Networks for One-shot Image Recognition. *ICML*.

[3] Vinyals O. et al. (2016) Matching Networks for One Shot Learning. *NIPS*.

[4] Snell J. et al. (2017) Prototypical Networks for Few-shot Learning. *NIPS*.

[5] Ren M et al. (2018) Meta-learning for semi-supervised few-shot classification. *ICLR*.

[6] Finn C. et al. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. arXiv:1703.03400. *ICML*.

[7] Ravi S. and Larochelle H. (2017) Optimization as a Model for Few-Shot Learning. *ICLR*.

[8] Li Z. et al. (2017) Meta-SGD: Learning to learn quickly for few shot learning. *arXiv*:1707.09835

[9] Mishra N. et al. (2018) A simple neural attentive meta-learner. *ICLR*.

[10] Lake B. M. et al. (2015). Human-level concept learning through probabilistic program induction. *Science*. Omniglot dataset: https://github.com/brendenlake/omniglot

[11] Guo Y. and Zhang L. (2017) One-shot Face Recognition by Promoting Underrepresented Classes. MS-Celeb-1M Challenge: Low-shot Learning. Task Description: Know you at One Glance. https://www.msceleb.org/challenge2/2017

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 : - Machine learning, deep learning, meta learning, vision par ordinateur - Tensorflow, PyTorch, python	