



CEA List Laboratoire d'Analyse Sémantique Texte-Image Centre de Saclay 91191 Gif-sur-Yvette France http://www.kalisteo.eu	Contact	
	Tél	+33 (0)1 69 08 01 54
	E-mail	adrian.popescu@cea.fr

STAGE LASTI-2021-S1

Apprentissage incrémental profond réaliste

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 d'Analyse Sémantique des Textes et des Images (LASTI) est une équipe de 25 personnes (chercheurs, ingénieurs, doctorants) menant des recherches sur les technologies de description et de compréhension du contenu multimédia (image, texte, parole) et des documents multilingues, en particulier à grande échelle. Les enjeux scientifiques sont :

- développer des algorithmes efficaces et robustes pour l'analyse et l'extraction de contenu multimédia, leur classification et analyse sémantique;
- reconstituer ou fusionner de données hétérogènes pour interpréter des scènes ou documents;
- développer des méthodes et des outils pour la construction, la formalisation et l'organisation des ressources et connaissances nécessaires au fonctionnement de ces algorithmes;
- intégrer plusieurs de ces briques technologiques afin d'accéder à l'information et répondre à un besoin utilisateur (moteurs de recherche, chatbot, rapports synthétiques de veille).

Contexte du stage

L'apprentissage incrémental est une branche de l'intelligence artificielle qui gagne en popularité de par son potentiel applicatif important. Son objectif global est d'apprendre continuellement à partir d'un flux de données dynamiques sans accès ou avec un accès limité aux données passées. Etant donné un système qui reconnaît un certain nombre de classes, il faut faire évoluer ce système vers un nouvel état afin d'incorporer de nouvelles classes au fur et à mesure que les données arrivent. L'apprentissage incrémental est intéressant lorsque les ressources en mémoire et en temps d'exécution sont limitées. Le défi majeur que rencontrent les systèmes d'apprentissages incrémental actuels est l'oubli catastrophique (catastrophic forgetting). Celui-ci traduit la tendance des réseaux de neurones à oublier les anciennes classes lorsque de nouvelles classes sont apprises.

Description du stage

Le stage se focalise la classification des images dans un contexte incrémental en utilisant les réseaux de neurones convolutifs. Les principaux objectifs seront de :

- étudier les méthodes d'apprentissage incrémental existantes et en proposer des méthodes qui passent à l'échelle en nombre de classes et de temps nécessaire pour leur apprentissage.
- améliorer la plasticité des algorithmes incrémentaux afin d'inclure des connaissances diversifiées lors des différentes phases de l'apprentissage en un seul modèle profond.

CEA List Laboratoire d'Analyse Sémantique Texte-Image Centre de Saclay 91191 Gif-sur-Yvette France http://www.kalisteo.eu	Contact	
	Tél	+33 (0)1 69 08 01 54
	E-mail	adrian.popescu@cea.fr

- mettre en place un cadre d'évaluation qui soit plus conforme aux exigences applicatives que les protocoles actuels. Il s'agira notamment de : varier le nombre de classes nouvelles par état incrémental, gérer des données bruitées, avoir une quantité de données variable par classe.

La mise en œuvre de ces objectifs s'appuiera sur des recherches en cours au laboratoire. Le poids donné à chaque objectif sera adapté en fonction des compétences et souhaits de la candidate ou du candidat sélectionné. Le stage nécessite des compétences en apprentissage profond qui seront développées pendant sa réalisation.

Selon les résultats obtenus, le stage pourra donner lieu à une publication scientifique.

Références

- E Belouadah, A Popescu . "IL2M: Class Incremental Learning with Dual Memory", Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2019, p583-592
- E Belouadah, A Popescu. « ScalL: Classifier Weights Scaling for Class Incremental Learnin », The IEEE Winter Conference on Applications of Computer Vision (WACV), 2020, p1266-1275
- E Belouadah, A Popescu, I Kanellos. « Initial Classifier Weights Replay for Memoryless Class Incremental Learning », British Machine Vision Conference (BMVC) 2020.
- Yue Wu, Yinpeng Chen, Lijuan Wang, Yuancheng Ye, Zicheng Liu, Yandong Guo, and Yun Fu. « Large scale incremental learning ». In IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019, pages 374–382, 2019.
- Saihui Hou, Xinyu Pan, Chen Change Loy, Zilei Wang, and Dahua Lin. « Learning a unified classifier incrementally via rebalancing ». In IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019, pages 831–839, 2019.
- Francisco M. Castro, Manuel J. Marín-Jiménez, Nicolás Guil, Cordelia Schmid, and Karteek Alahari. « End-to-end incremental learning ». In ECCV 2018- 15th European Conference in Computer Vision, Munich, Germany, September 8-14, 2018, Proceedings, Part XII, pages 241–257, 2018.
- Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H. Lampert. « icarl: Incremental classifier and representation learning ». In Conference on Computer Vision and Pattern Recognition, CVPR, 2017

Niveau demandé :	Ingénieur, Master 2
Ce stage ouvre la possibilité de poursuite en thèse ou 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"> - Apprentissage profond (deep learning) - Vision par ordinateur - Python 	



CEA List Laboratoire d'Analyse Sémantique Texte-Image Centre de Saclay 91191 Gif-sur-Yvette France http://www.kalisteo.eu	Contact	
	Tél	+33 (0)1 69 08 01 54
	E-mail	adrian.popescu@cea.fr

- Pytorch (requis)
- Tensorflow (pas requis, mais préférable)