Theses 2022   Ref : LVA-22-T3

Continual Learning for Scene Analysis in Open World

Presentation of the host laboratory

Based in Paris-Saclay campus, CEA-LIST is one of four technological research institutes of CEA TECH, the technological research direction of CEA. Dedicated to intelligent digital systems, it contributes to the competitiveness of companies via research and knowledge transfers. The expertise and competences of the 800 research engineers and technicians at CEA-LIST help more than 200 companies in France and abroad every year on subjects categorized over 4 programs and 9 technological platforms. 21 start-ups have been created since 2003.

The Computer Vision and Machine Learning for scene understanding laboratory addresses computer vision subjects with a stronger emphasis on four axes:

- Recognition (detection or segmentation of objects and persons)
- Behavior analysis (action and gesture recognition, anomalous behavior of individuals or crowds)
- Smart annotation (large scale annotation of 2D and 3D data using semi-supervised methods)
- Perception and decision-making (Markovian decision processes, navigation)

The PhD candidate will join a team composed of 30 researchers (researchers, PhD students, interns) and will be able to interact with peers working on related subjects and methods.

Context

Visual scene understanding is necessary in many applications. To this aim, deep learning models are particularly efficient to deal with computer vision tasks like semantic segmentation and object detection. However, a trained model is no more effective nor useful when seeing new semantic classes ("things" or "stuff") which were not present during the training. Training from scratch the model with additional labeled images representing this novelty together with the old images would not be efficient or possible (too large size or unavailability of old data).

Incrementally adapting an existing object detection model to detect new unseen classes with severe memory and computational constraints is a critical capacity in real-word applications such as robotics, self-driving vehicles or video surveillance. However, while human beings can easily recognize new objects continuously without forgetting the old knowledge, deep learning models can suffer from 'catastrophic forgetting'. In fact, adding new classes without using the old training dataset can cause a big degradation of performance on the original set of classes.

To overcome this issue, many approaches of continual/incremental learning have been proposed [1]. Some methods use a memory buffer to save a set of the old dataset and re-use it to retrain the model with the new classes [2] or extend the model architecture by adding other detection heads. Others focus essentially on regularizing the training to minimize the discrepancy between responses for the old and the updated model [3]. These methods have to face a trade-off between model plasticity and rigidity. Their results are still limited compared to the models trained jointly with all the dataset.

While most of studies are conducted on image classification [4], fewer methods focus on incremental learning for tasks like object detection or semantic segmentation [5]. In general, models learn a little from new classes and try not to degrade performance on old classes whereas it would be desirable to learn better new classes and improve old classes that can also be present in new data. Moreover, adapting to these tasks methods of incremental learning developed for image classification is not straightforward. Indeed, the additional challenge of 'background shift' makes the incremental learning more difficult.

To this aim, recent methods make the open-set hypothesis to detect a novelty, out of the known classes. Some methods go further by differentiating unknown classes of objects, with the open-world hypothesis [6, 7, 8, 9, 10]. They identify instances of unknown objects as unknown and subsequently learn to recognize them when training data progressively arrive without retraining from scratch (cf. Fig.1). Similarly, methods for open-world semantic segmentation have been proposed [11, 12].
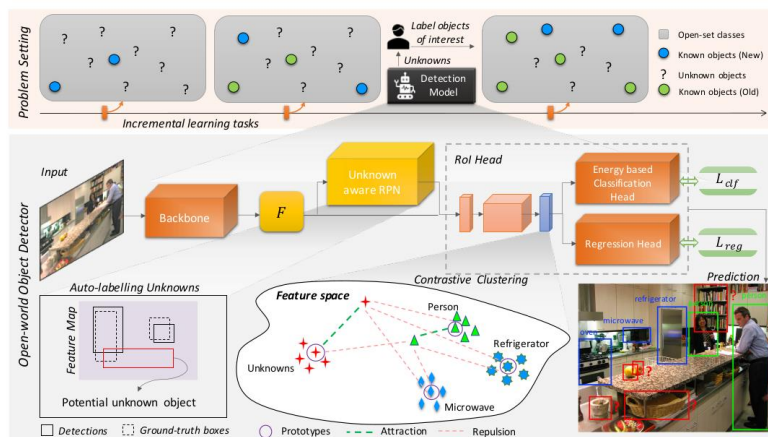
Figure 2: *Approach Overview: Top row:* At each incremental learning step, the model identifies unknown objects (denoted by '?'), which are progressively labelled (as blue circles) and added to the existing knowledge base (green circles). *Bottom row:* Our open world object detection model identifies potential unknown objects using an energy-based classification head and the unknown-aware RPN. Further, we perform contrastive learning in the feature space to learn discriminative clusters and can flexibly add new classes in a continual manner without forgetting the previous classes.

Figure 1 (extracted from [7]): An example of incremental learning approach for object detection

Objective of the thesis

The objective of the thesis is to propose methods of continual learning with the open-world hypothesis in order to give to scene understanding models (like semantic segmentation and object detection) the ability of evolving when seeing new semantic classes. This evolution must be made with the constraint of unavailability of past data (hypothesis of incremental learning).

The ability of discovering unseen classes and recognizing them as a novelty will be necessary in order to make the model effective and trustable. The ability of distinguishing different unseen classes and different instances of unseen classes will also be advantageous for a more efficient learning. In addition, we will study how the few-shot learning paradigm can be applied efficiently to learn the new classes with fewer annotations [13, 14], therefore, minimizing human effort.

As usual, the PhD candidate will start with analyzing the state of the art in open-world incremental learning and pointing the limitations of the existing methods. Novel methods with severe memory, annotation and computational constraints will be proposed, developed, evaluated with appropriate benchmarks then published.

References

[1] Parisi et al. Continual Lifelong Learning with Neural Networks: A Review. Neural Networks, Feb. 2019.

[2] Shmelkov et al. Incremental Learning of Object Detectors without Catastrophic Forgetting, *ICCV* 2017.

[3] Shieh et al. Continual Learning Strategy in One-Stage Object Detection Framework Based on Experience Replay for Autonomous Driving Vehicle, *Sensors* 20, no. 23: 6777. 2020.

[4] De Lange et al. A continual learning survey: Defying forgetting in classification tasks. IEEE Tr. PAMI 2022.

[5] Michieli et al. Knowledge Distillation for Incremental Learning in Semantic Segmentation. CVIU 2021.

[6] Singh et al. ORDER: Open World Object Detection on Road Scenes. Machine Learning for Autonomous Driving Workshop, NeurIPS 2021.

[7] Joseph et al. Towards Open World Object Detection. CVPR 2021.

[8] Gupta et al. OW-DETR: Open-world Detection Transformer. CVPR 2022.

[9] Kim et al. Learning Open-World Object Proposals Without Learning to Classify. IEEE Robotics and Automation Letters, April 2022.

[10] Konan et al. Extending One-Stage Detection with Open-World Proposals. arXiv:2201.02302

[11] Cen et al. Deep Metric Learning for Open World Semantic Segmentation. ICCV 2021.

[12] Hong et al. GOSS: Towards Generalized Open-set Semantic Segmentation. arXiv:2203.12116

[13] Li et al. Towards Generalized and Incremental Few-Shot Object Detection. arXiv:2109.11336

[14] Pérez-Rua et al. Incremental Few-Shot Object Detection. CVPR 2020.

Keywords

Continual learning, open-world incremental learning, knowledge distillation, object detection, semantic segmentation, deep learning.

Required profile

- Research master or engineer with experience in computer vision and machine learning
- Skills: computer vision, machine learning, including deep learning
- Programming: Python, Deep Learning frameworks (Tensorflow, PyTorch)

If you recognize yourself in these skills and this profile and if developing your skills on the theme of AI, within an ambitious and recognized institute, at the heart of the dynamic environment of the Saclay plateau is part of your professional project, please send CV + cover letter to:

romaric.audigier@cea.fr / hejer.ammar@cea.fr