Hierarchical Learning for Better Concept Generalization
– Master 2 internship 2023 –

Application deadline : December 10th, 2022

Location	Cnam, Paris (France)
Salary	≈ 600€/month
Contract	Internship (4 to 6 months)
Starting date	Flexible in 2023

Context

Deep neural networks (DNN) have become very successful in computer vision, in tasks such as image classification [5]. As the cost of training DNNs can be cumbersome, a new sought-after property of DNNs is *concept generalization* [8]. When trained on some concepts, *e.g.* "dogs" and "cats", a model presenting good generalization properties, is able to learn with *few* examples semantically close concepts, such as "wolves", but also very different concepts, such as "cars" or "plane". The recent Imagenet-CoG benchmark [8], illustrated in Fig. 1, has been introduced to measure this property of DNNs. A model is evaluated by first being trained on the 1000 classes of Imagenet-1k [3] (pink on Fig. 1). Then its features are used to classify *unseen* and progressively more challenging subsets of concepts (orange to bright yellow on Fig. 1). The final CoG score is the overall accuracy on each subsets.

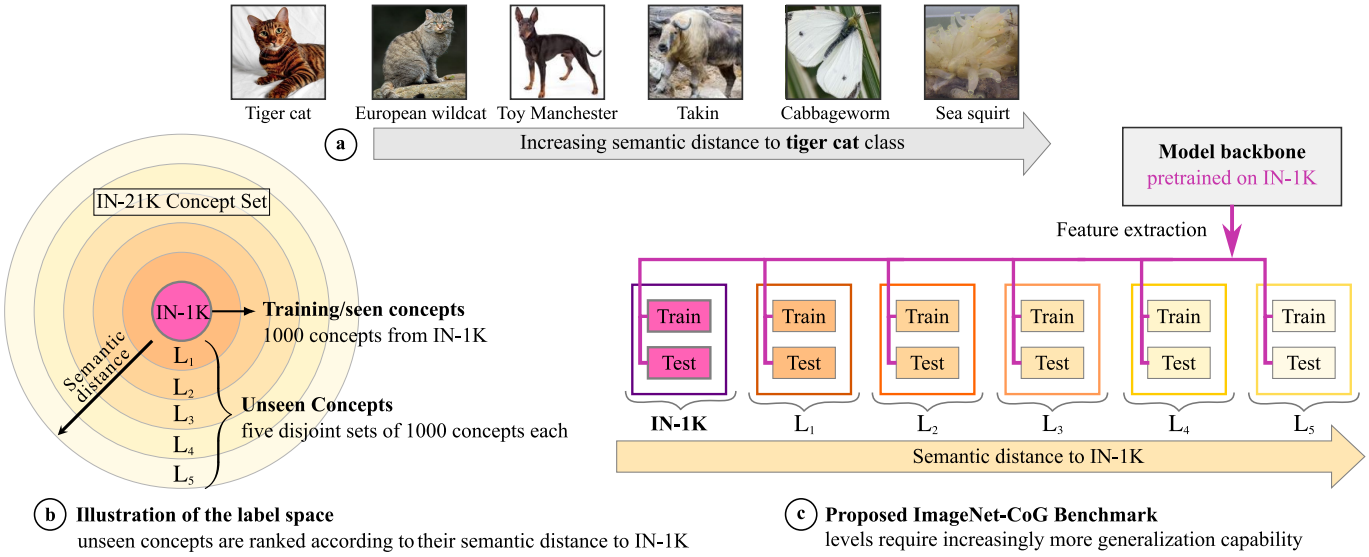

FIGURE 1 – The Imagenet-Cog benchmark. Figure taken from [8].

Concept generalization is a hard task : not only deep models must exhibit strong performances on the original task they were trained for (*e.g.* image classification on Imagenet), but they should also be able to generalize with satisfactory performances on yet unknown tasks. This property is of strong interest to the artificial intelligence community as it is linked to applications in lifelong learning, domain adaptation, and open-set generalization. Promising first results have been obtained using unsupervised learning, *e.g.* [2], these methods lead to stronger generalization [8] than standard supervised models when measured on the CoG benchmark. Recently dedicated supervised methods have been introduced, *e.g.* [7], to train models with better CoG scores.

It was noted in [4] that learning the underlying hierarchical structure of labels can yield a better model for "downstream" tasks, *i.e.* tasks that the model was not trained for (*e.g.* object detection, image segmentation). In [6] models trained while considering hierarchical relations between labels have a more structured image embedding space. It was not studied however if models leveraging hierarchical labels structure lead to better CoG scores.

Internship objectives

The goal of the internship is to investigate whether concept generalization can benefit from the introduction of prior hierarchical knowledge into the training process of deep neural networks (DNN). More precisely, we aim to leverage hierarchical relations between labels during the training of DNNs in order to improve the concept generalization on the ImageNet-CoG benchmark [8].

During the internship the student will :

1. Review the state of the art on topics such as generalization and hierarchical training;
2. Use publicly available code to evaluate baseline models on the CoG benchmark [1];
3. Train an improved model leveraging a hierarchical loss, *e.g.* [1], and evaluate the expected improvements on CoG over the baseline;
4. If there is time, investigate if the same conclusion holds for "hierarchical" unsupervised methods, such as [4].

The student will be using new training schemes [9] to ease the computational costs of DNN training on the large Imagenet [3] database (1m images).

Applicant profile

The ideal applicant is pursuing an MSc. in Computer Science, Artificial Intelligence or Computer Vision or an equivalent degree (*e.g.* engineering diploma). A good grasp of the fundamentals of machine learning and deep learning is expected. The candidate should have an interest for scientific research and good written and oral communication skills. Knowledge of the Python programming language is a must, including some level of experience with at least one deep learning framework (PyTorch, Keras, TensorFlow, JAX…). A first experience with computer vision, image processing or unsupervised learning is a plus. All applications, independently of previous experience, will be considered, provided that the candidate motivation and profile fits the internship topic. French is not required but can help with the everyday life.

Where you will work

This position is for an internship from 4 to 6 months, with a flexible starting date in 2023. The internship will take place at the *Center for research and studies in computer science and communications* (CEDRIC), the computer science laboratory of the Conservatoire national des arts et métiers (Cnam), a prestigious French higher education institute. It is located in Paris (2 rue Conté, 3rd "arrondissement"). The intern will join the *Complex Data, Machine Learning and Representations* team [2] and will be supervised by Nicolas Audebert, Clément Rambour and Elias Ramzi.

How to apply

Please send your application (CV and a short motivation letter) to `elias.ramzi@lecnam.net`, `nicolas.audebert@cnam.fr` and `clement.rambour@cnam.fr`.

1. https://github.com/naver/cog
2. https://cedric.cnam.fr/lab/equipes/vertigo/

Références

[1] Luca Bertinetto, Romain Mueller, Konstantinos Tertikas, Sina Samangooei, and Nicholas A Lord. Making better mistakes : Leveraging class hierarchies with deep networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12506–12515, 2020.

[2] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9650–9660, 2021.

[3] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet : A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[4] Yuanfan Guo, Minghao Xu, Jiawen Li, Bingbing Ni, Xuanyu Zhu, Zhenbang Sun, and Yi Xu. Hcsc : Hierarchical contrastive selective coding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9706–9715, 2022.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[6] Elias Ramzi, Nicolas Audebert, Nicolas Thome, Clément Rambour, and Xavier Bitot. Hierarchical average precision training for pertinent image retrieval. In *Proceedings of the IEEE/CVF European Conference on Computer Vision*, 2022.

[7] Mert Bulent Sariyildiz, Yannis Kalantidis, Karteek Alahari, and Diane Larlus. Improving the generalization of supervised models. *arXiv preprint arXiv :2206.15369*, 2022.

[8] Mert Bulent Sariyildiz, Yannis Kalantidis, Diane Larlus, and Karteek Alahari. Concept generalization in visual representation learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9629–9639, 2021.

[9] Ross Wightman, Hugo Touvron, and Hervé Jégou. Resnet strikes back : An improved training procedure in timm. *arXiv preprint arXiv :2110.00476*, 2021.