Three types of incremental learning

Gido M. van de Ven[1,2,3,*][0000−0002−5239−5660], Tinne Tuytelaars[3][0000−0003−3307−9723], and Andreas S. Tolias[1,4][0000−0002−4305−6376]

[1] Center for Neuroscience and AI, Baylor College of Medicine, Houston, US
[2] Computational and Biological Learning Lab, University of Cambridge, UK
[3] Department of Electrical Engineering, KU Leuven, Belgium
[4] Department of Electrical and Computer Engineering, Rice University, Houston, US
[*] gido.vandeven@kuleuven.be

Abstract. This is a "Type B: Encore abstracts" submission based on an article published in *Nature Machine Intelligence* (date of acceptance: 18 October 2022): https://doi.org/10.1038/s42256-022-00568-3.

Keywords: Continual learning · Lifelong learning · Incremental learning

An important open problem in deep learning is enabling neural networks to incrementally learn from non-stationary streams of data [1,3]. For example, when deep neural networks are trained on samples from a new task or data distribution, they tend to rapidly lose previously acquired capabilities, a phenomenon referred to as catastrophic forgetting [7,2]. In stark contrast, humans and other animals are able to incrementally learn new skills without compromising those that were already learned [4]. The field of continual learning, also referred to as lifelong learning, is devoted to closing the gap in incremental learning ability between natural and artificial intelligence. In recent years, this area of machine learning research has been rapidly expanding, fuelled by the potential utility of deploying continual learning algorithms for applications such as medical diagnosis [5], autonomous driving [10] or predicting financial markets [9].

Despite its scope, continual learning research is relatively unstructured and the field lacks a shared framework. Because of an abundance of subtle, but often important, differences between evaluation protocols, systematic comparison between continual learning algorithms is challenging, even when papers use the same datasets [8]. It is therefore not surprising that numerous continual learning methods claim to be state-of-the-art. To help address this, here we describe a structured and intuitive framework for continual learning.

We put forward the view that, at the computational level [6], there are three fundamental types, or 'scenarios', of supervised continual learning (Table 1). Informally, (a) in task-incremental learning, an algorithm must incrementally learn a set of clearly distinguishable tasks; (b) in domain-incremental learning, an algorithm must learn the same kind of problem but in different contexts; and (c) in class-incremental learning, an algorithm must incrementally learn to distinguish

Table 1. — Overview of the three continual learning scenarios. Notation: $\mathcal{X}$ is the input space, $\mathcal{Y}$ is the within-context output space and $\mathcal{C}$ is the context space. In this article, the term 'context' refers to an underlying distribution from which observations are sampled. The context changes over time. In the continual learning literature, the term 'task' is often used in a way analogous to how the term 'context' is used here. Task-IL: task-incremental learning, Domain-IL: domain-incremental learning, Class-IL: class-incremental learning.

Scenario	*Intuitive description*	*Mapping to learn*
Task-IL	Sequentially learn to solve number of distinct tasks	$f\colon \mathcal{X} \times \mathcal{C} \to \mathcal{Y}$
Domain-IL	Learn to solve same problem in different contexts	$f\colon \mathcal{X} \to \mathcal{Y}$
Class-IL	Discriminate between incrementally observed classes	$f\colon \mathcal{X} \to \mathcal{C} \times \mathcal{Y}$

between a growing number of objects or classes. In the accompanying article, we formally define these three scenarios and point out different challenges associated with each one of them. We also review existing strategies for continual learning with deep neural networks and we provide a comprehensive, empirical comparison to test how suitable these different strategies are for each scenario.

These three scenarios and their different challenges can be conveniently studied in an academic continual learning setting, where a classification-based problem is split up in discrete, non-overlapping contexts (which are often called 'tasks') that are encountered in sequence. We show that in this setting there is a clear separation between the three scenarios. At least in part because of a preprint of the accompanying article [11], the terms 'task-incremental learning', 'domain-incremental learning' and 'class-incremental learning' are sometimes used in the recent literature in a way that restricts them to this academic setting. Here, by interpreting these three scenarios as specifying how the non-stationary aspect of the data relates to the mapping that must be learned, we propose that they generalize to more flexible continual-learning settings. To demonstrate the value of such generalized versions of these three scenarios, we show how a 'task-free' data stream without sharp context boundaries can also be performed in these three different ways.

We believe that the three continual learning scenarios described in this article provide a useful basis for defining clear and unambiguous benchmark problems for continual learning. We hope this will accelerate progress to bridge the gap between natural and artificial intelligence. Moreover, we believe that it is an important conceptual insight that, at the computational level, a supervised learning problem can be incremental in these three different ways. Perhaps especially in the real world, where continual learning problems are often complex and 'mixtures' of scenarios, it might be fruitful to approach problems as consisting of a combination of these three fundamental types of incremental learning.

References

1. Chen, Z., Liu, B.: Lifelong machine learning. Synthesis Lectures on Artificial Intelligence and Machine Learning **12**(3), 1–207 (2018)
2. French, R.M.: Catastrophic forgetting in connectionist networks. Trends in cognitive sciences **3**(4), 128–135 (1999)
3. Hadsell, R., Rao, D., Rusu, A.A., Pascanu, R.: Embracing change: Continual learning in deep neural networks. Trends in Cognitive Sciences **24**(12), 1028–1040 (2020)
4. Kudithipudi, D., Aguilar-Simon, M., Babb, J., Bazhenov, M., Blackiston, D., Bongard, J., Brna, A.P., Raja, S.C., Cheney, N., Clune, J., Daram, A., Fusi, S., Helfer, P., Kay, L., Ketz, N., Kira, Z., Kolouri, S., Krichmar, J.L., Kriegman, S., Levin, M., Madireddy, S., Manicka, S., Marjaninejad, A., McNaughton, B., Miikkulainen, R., Navratilova, Z., Pandit, T., Parker, A., Pilly, P.K., Risi, S., Sejnowski, T.J., Soltoggio, A., Soures, N., Tolias, A.S., Urbina-Meléndez, D., Valero-Cuevas, F.J., van de Ven, G.M., Vogelstein, J.T., Wang, F., Weiss, R., Yanguas-Gil, A., Zou, X., Siegelmann, H.: Biological underpinnings for lifelong learning machines. Nature Machine Intelligence **4**(3), 196–210 (2022)
5. Lee, C.S., Lee, A.Y.: Clinical applications of continual learning machine learning. The Lancet Digital Health **2**(6), e279–e281 (2020)
6. Marr, D.: Vision: A computational investigation into the human representation and processing of visual information. W.H. Freeman and Company (1982)
7. McCloskey, M., Cohen, N.J.: Catastrophic interference in connectionist networks: The sequential learning problem. In: Psychology of learning and motivation, vol. 24, pp. 109–165. Elsevier (1989)
8. Mundt, M., Lang, S., Delfosse, Q., Kersting, K.: CLEVA-compass: A continual learning evaluation assessment compass to promote research transparency and comparability. In: International Conference on Learning Representations (2022)
9. Philps, D., Weyde, T., Garcez, A.d., Batchelor, R.: Continual learning augmented investment decisions. Preprint at `https://arxiv.org/abs/1812.02340` (2018)
10. Shaheen, K., Hanif, M.A., Hasan, O., Shafique, M.: Continual learning for real-world autonomous systems: Algorithms, challenges and frameworks. Journal of Intelligent & Robotic Systems **105**(1), 1–32 (2022)
11. van de Ven, G.M., Tolias, A.S.: Three scenarios for continual learning. Preprint at `https://arxiv.org/abs/1904.07734` (2019)