Brain-Inspired Replay in Artificial Neural Networks

Gido M. van de Ven[1,2] & Andreas S. Tolias[1,3]

[1] Center for Neuroscience and Artificial Intelligence, Department of Neuroscience, Baylor College of Medicine, Houston, US
[2] Computational and Biological Learning Lab, Department of Engineering, University of Cambridge, UK
[3] Department of Electrical and Computer Engineering, Rice University, Houston, US

Introduction

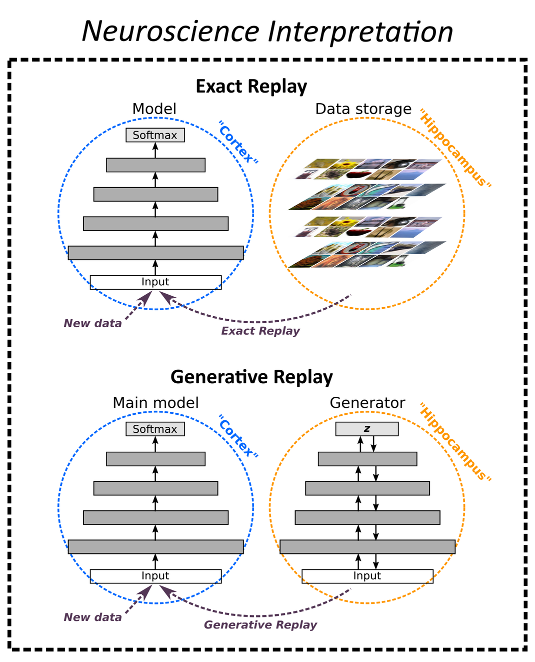
State-of-the-art deep neural networks can solve about any problem they are trained on. But when trained on a new problem, these networks quickly forget any previously learned problems. A simple solution for such 'catastrophic forgetting' is to store the encountered examples from previously learned problems and revisit them when learning about new problems. Despite being effective, such ‘replay' or ‘rehearsal' is typically believed not to be a scalable solution as (1) constantly retraining on all previous problems is considered very inefficient and (2) the amount of data that would have to be stored quickly becomes unmanageable. Yet, in the brain—which clearly has implemented an efficient and scalable algorithm for continual learning—the replay of previous experiences is important for the gradual stabilization of new memories [1,2]. Inspired by this, here we revisit the use of replay as a tool for continual learning in artificial neural networks (ANNs).

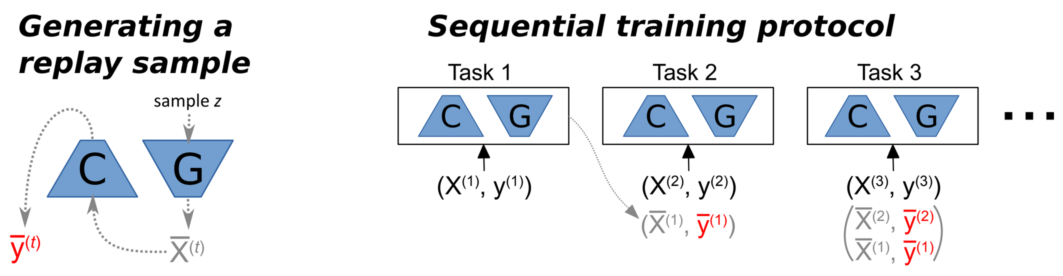
Part I: A Strong Case for Replay

Methods 1: Adding replay to artificial neural networks

Exact Replay (ER): Data from previous tasks is stored and interleaved with the training data of later tasks. An important disadvantage is that storing data is not always possible (e.g., privacy concerns, limited storage capacity).

Generative Replay (GR): An alternative is to generate the data to be replayed. Besides the main model (**M**) for solving the tasks (e.g., a classifier), a separate generator (**G**) is trained [3]. We use a standard variational auto-encoder (VAE). When training on a new task, images generated by the (previous) generator are labelled by the (previous) classifier and interleaved with the current task's training data.





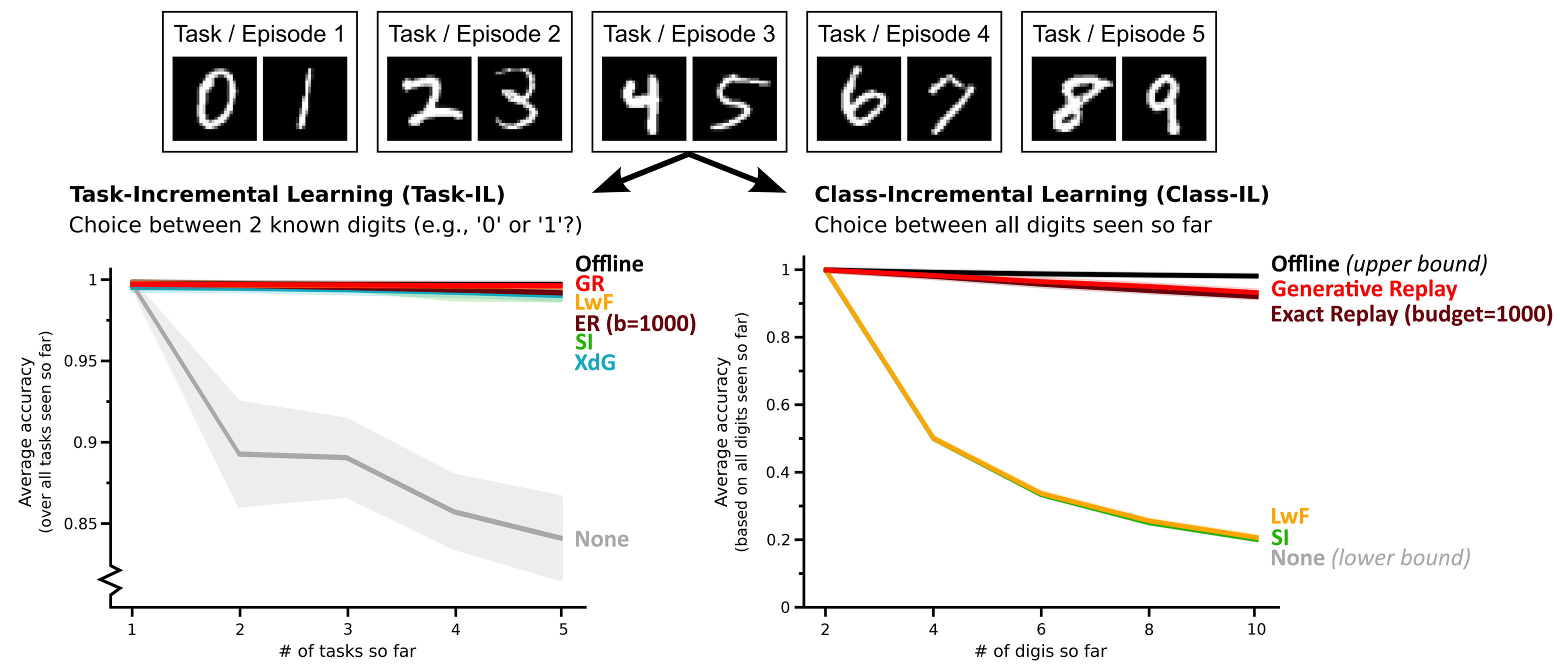
Methods 2: Other continual learning methods

Learning without Forgetting (LwF): Inputs from the current task are replayed, after they have been labelled by the model trained on the previous task [4].

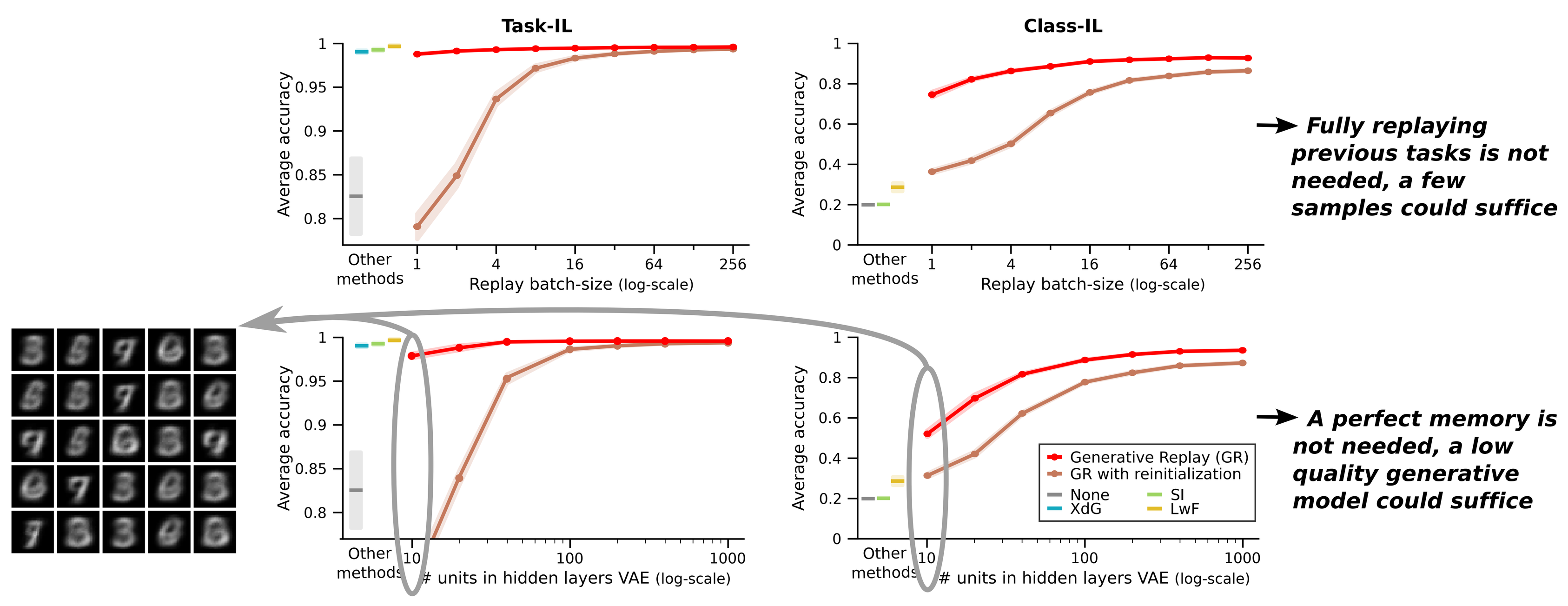
Synaptic Intelligence (SI): This metaplasticity-inspired method slows down learning for parameters that are important for previously learned task, by adding a regularization term to the loss that penalizes changes to parameters depending on their estimated importance [5].

Context-dependent Gating (XdG): For each task, a different random subset of units is gated (i.e., their activations set to zero) [6]. As this method requires availability of task identity at test time, it cannot always be used.

Results 1: Class-incremental learning *requires* replay

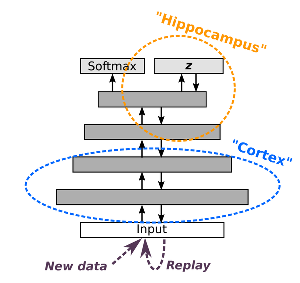


Results 2: Replay is surprisingly efficient and robust



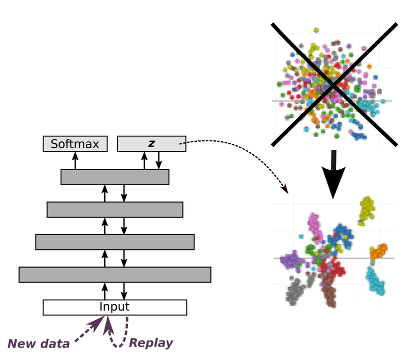
Part II: Brain-Inspired Modifications Enable Generative Replay to Scale

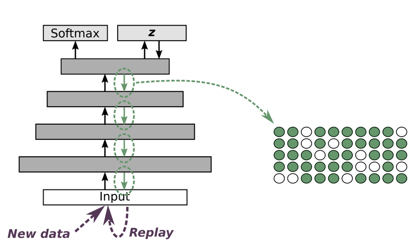
Methods 3: Brain-inspired replay (BI-R)



Replay-through-Feedback: Merge the generator into the main model. Rreplay is now generated by the network's feedback or backward connections.

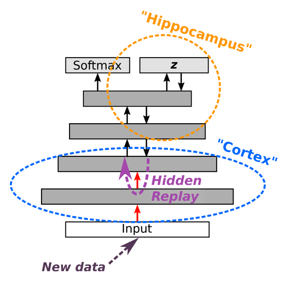
Conditional Replay: Enable the network to generate specific categories by replacing the standard normal prior by a Gaussian mixture with a separate mode for each class.



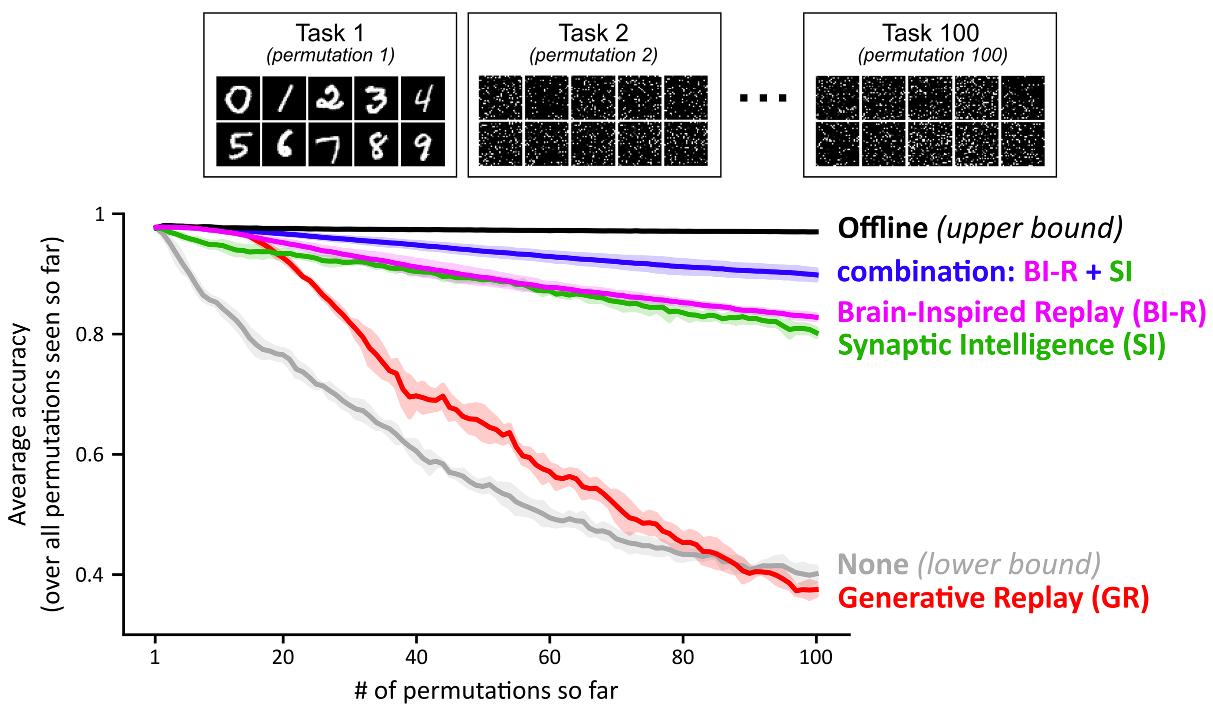


Context Gates: For each context or task, inhibit (or gate) a different, randomly selected subsets of neurons only during the feedback pass.

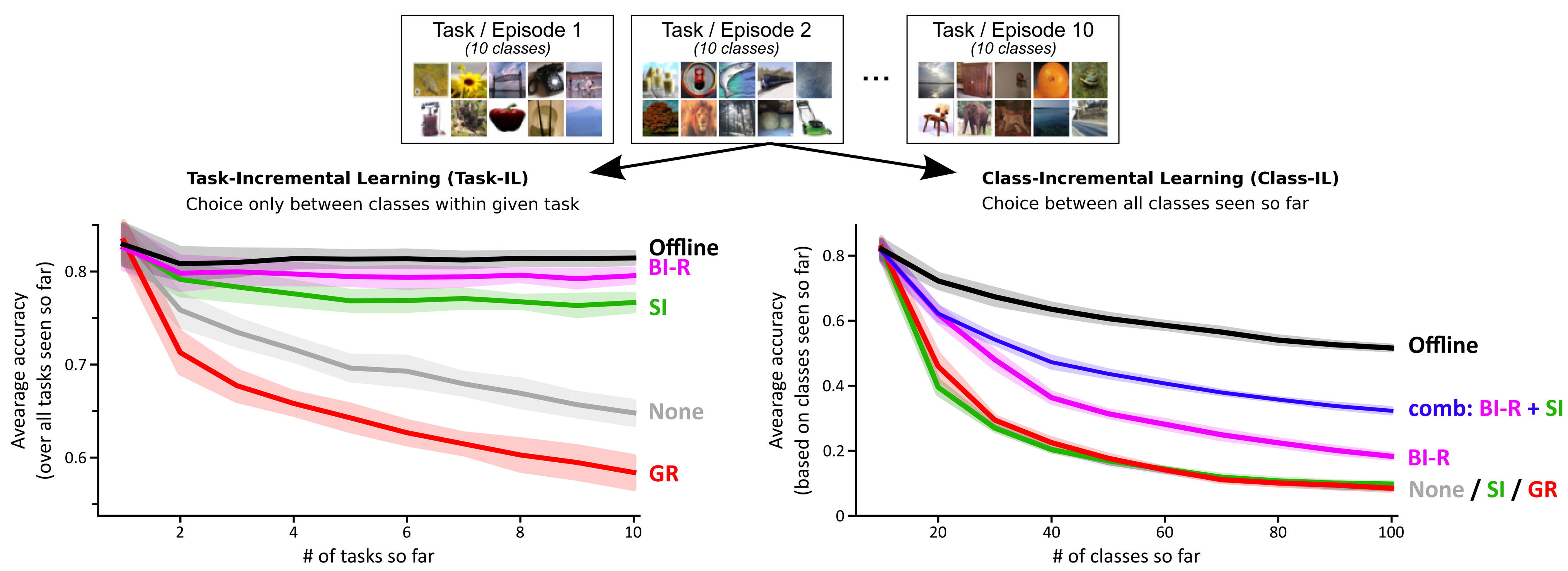
Hidden Replay: Replay internal representations, instead of representations at the input level (e.g., pixel level).



Results 3: Scaling up to many tasks



Results 4: Scaling up to complex inputs



Discussion

- Replay might be necessary for incremental class learning
- Generative replay can be surprisingly efficient and robust
- Easy-to-implement, brain-inspired modifications enable generative replay to scale to more challenging problems
- The "Brain-Inspired Replay framework" has potential as novel computational model for replay in the brain

References

[1] Wilson & McNaughton (1994) Reactivation of hippocampal ensemble memories during sleep. *Science* **265**: 676-679.
[2] van de Ven *et al.* (2016) Hippocampal offline reactivation consolidates recently formed cell assembly patterns during sharp wave-ripples. *Neuron* **92**: 968-974.
[3] Shin *et al.* (2017) Continual learning with deep generative replay. *NIPS*: 2994-3003.
[4] Li & Hoiem (2017) Learning without forgetting. *IEEE Trans Pattern Anal Mach Intell*.
[5] Zenke *et al.* (2017) Improved multitask learning through synaptic intelligence. *arXiv*: 1703.04200.
[6] Masse *et al.* (2018) Alleviating catastrophic forgetting using context-dependent gating and synaptic stabilization. *PNAS* **115**: E10467-E10475.

Acknowledgements

This research project has been supported by an IBRO-ISN Research Fellowship, by the Lifelong Learning Machines (L2M) program of the Defence Advanced Research Projects Agency (DARPA) via contract number HR0011-18-2-0025 and by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior/Interior Business Center (DoI/IBC) contract number D16PC00003. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, IARPA, DoI/IBC, or the U.S. Government.