KU LEUVEN

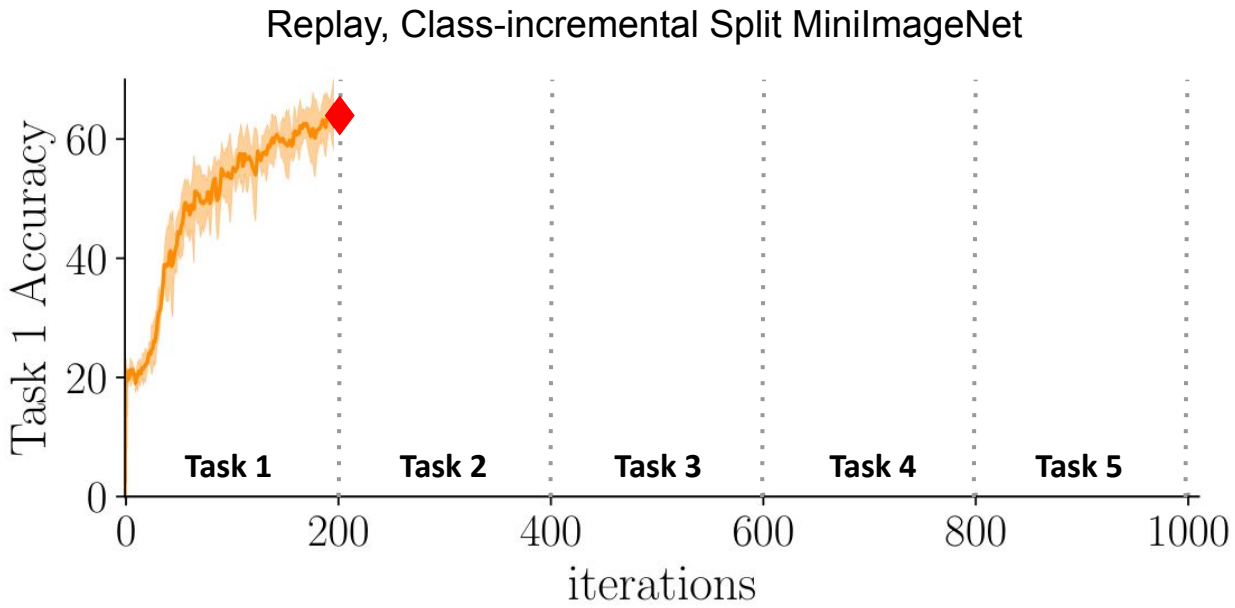
The “Stability Gap”

Gido van de Ven

Dagstuhl Seminar, 22 March 2023

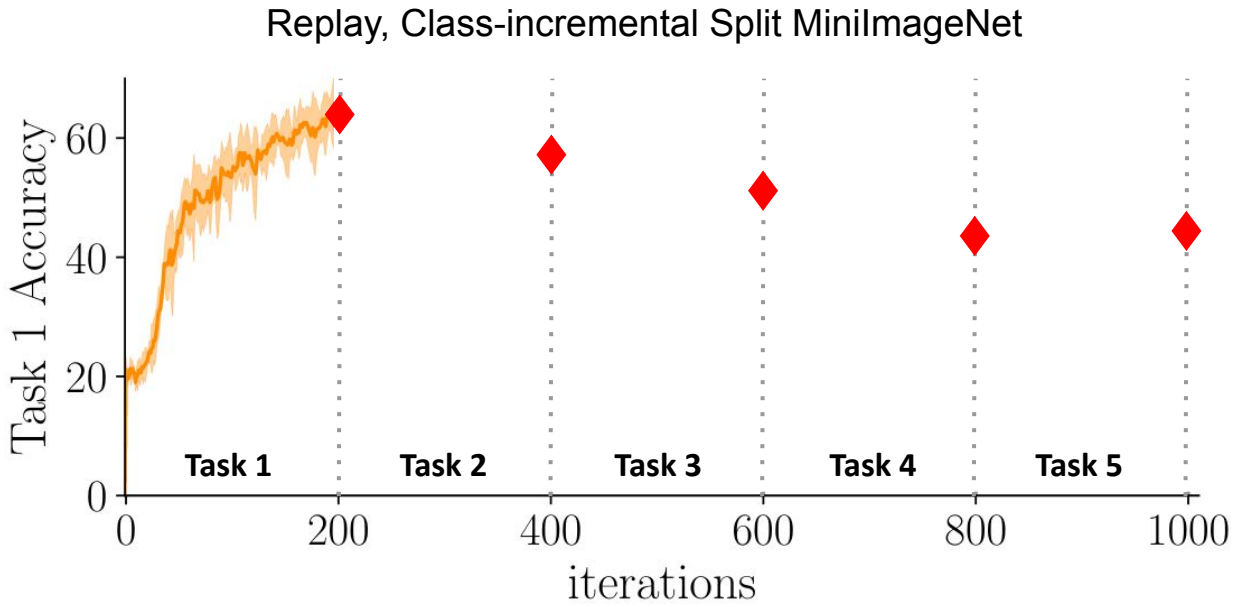
Based on work with Matthias De Lange & Tinne Tuytelaars

Does replay prevent forgetting?



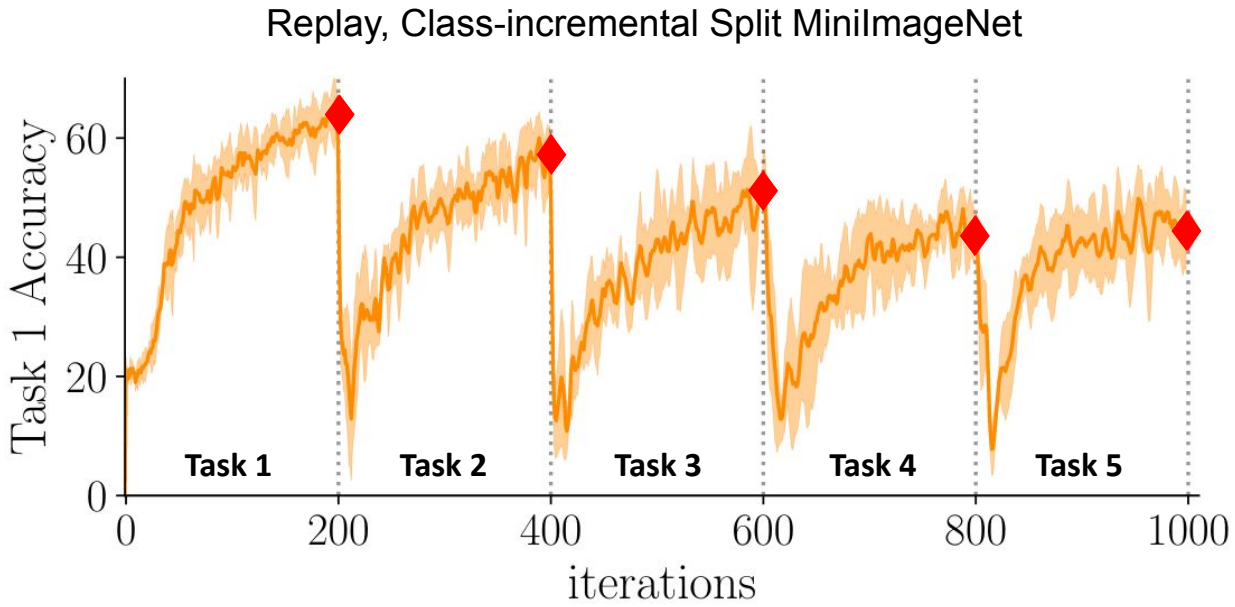
De Lange, van de Ven & Tuytelaars (2023) "Continual evaluation for lifelong learning: Identifying the stability gap" *ICLR, spotlight*

Does replay prevent forgetting?



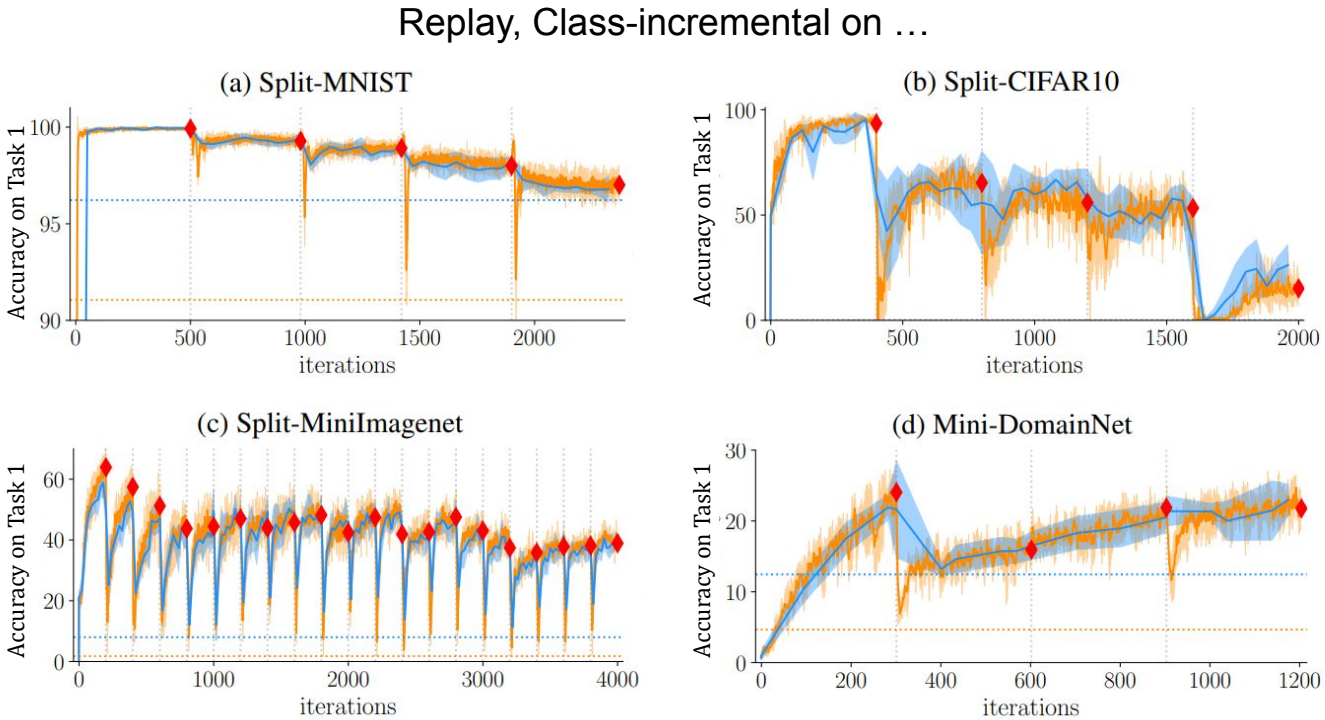
De Lange, van de Ven & Tuytelaars (2023) "Continual evaluation for lifelong learning: Identifying the stability gap" *ICLR, spotlight*

Does replay prevent forgetting?



De Lange, van de Ven & Tuytelaars (2023) "Continual evaluation for lifelong learning: Identifying the stability gap" *ICLR, spotlight*

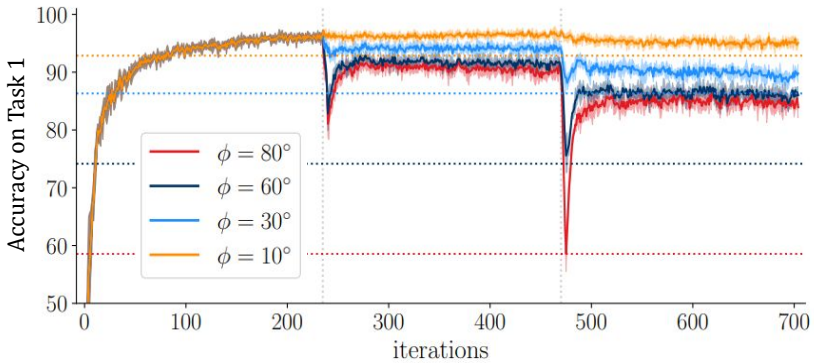
The stability gap is consistently observed



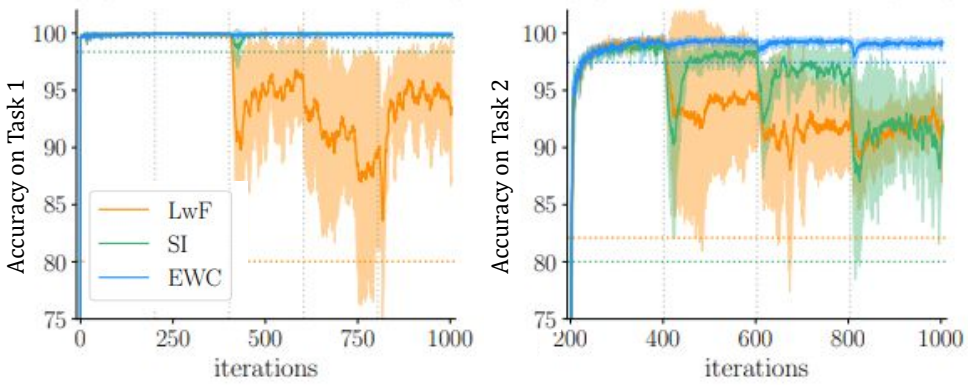
De Lange, van de Ven & Tuytelaars (2023) "Continual evaluation for lifelong learning: Identifying the stability gap" *ICLR, spotlight*

… also in other settings and methods

Replay, Domain-incremental Rotated MNIST



Regularization, Task-incremental Split MNIST



De Lange, van de Ven & Tuytelaars (2023) "Continual evaluation for lifelong learning: Identifying the stability gap" *ICLR, spotlight*

Why should we care?

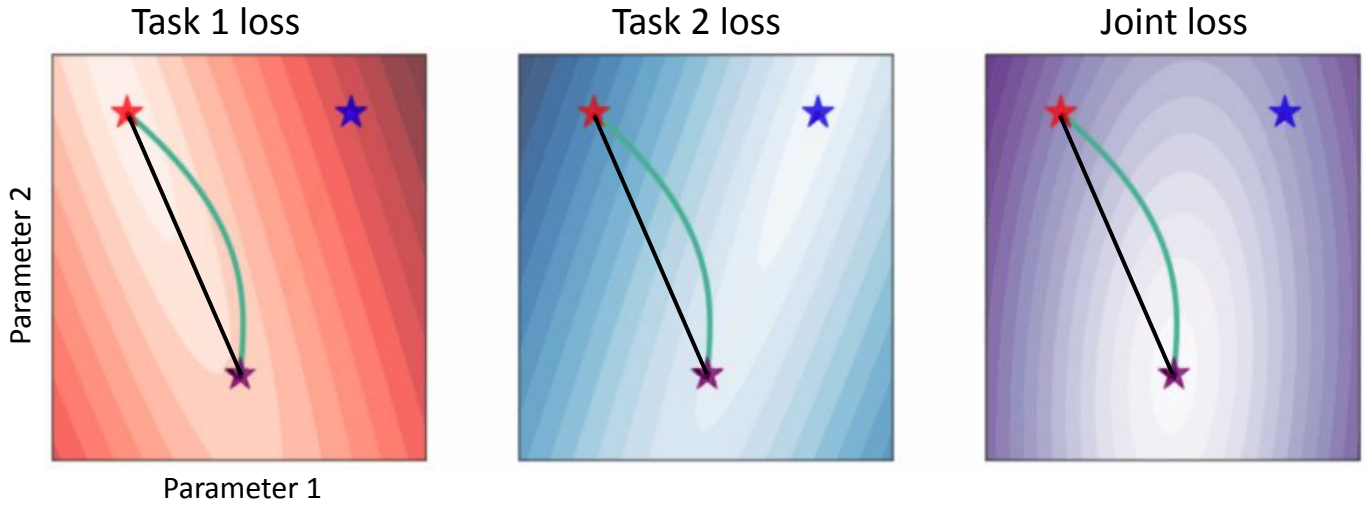
- Problematic for safety-critical applications
  - Worst-case performance might be important
  - Could be exploitable by adversarial agent with control over the training stream
- Could avoiding the stability gap lead to better final performance?
  - Preventing forgetting seems more efficient than having to re-learn
- Scientifically interesting
  - Insight into how replay works
  - Do humans suffer from transient forgetting upon learning something new?

Why does the stability gap happen?

- Quality of replay?
- Approximate nature of the SGD optimization? (e.g., tug-of-war dynamics)
- Specific to deep neural networks?

(also see the paper for a conceptual analysis with insights on this)

Stability gap in a toy example



Kao, Jensen et al. (2021) "Natural continual learning: success is a journey, not (just) a destination" *NeurIPS*

Questions?

Funding Acknowledgement

This project received funding from the ERC project KeepOnLearning (ref number 101021347), the KU Leuven C1 project Macchina, and the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 101067759.