On the Effectiveness of Lipschitz-Driven Rehearsal in Continual Learning – Additional illustrations

In this document, we enclose extra illustrations to complement the responses provided in the rebuttal.

Illustration of the Effect of LiDER – [PLZ4 - Additional Clarifications]

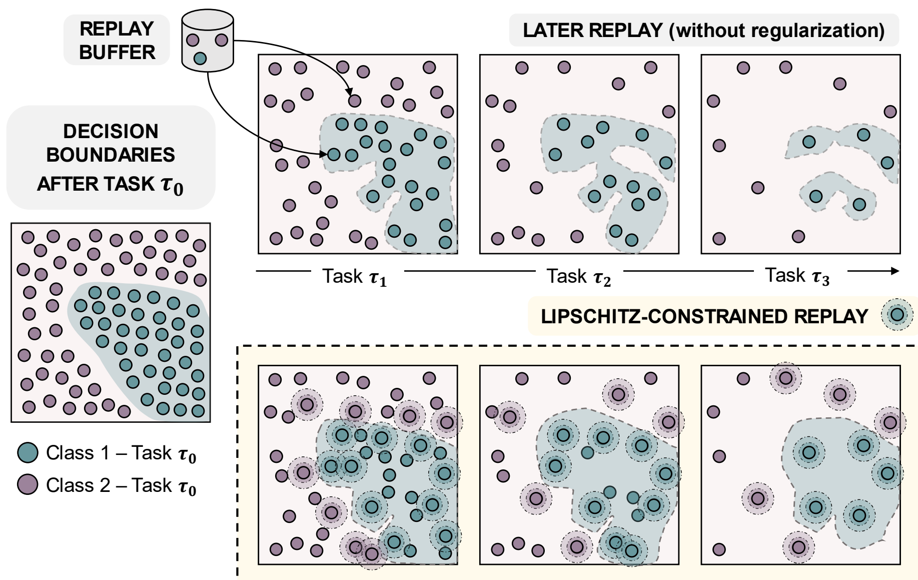

Figure A: Left: depiction of an initial decision boundary learned by the model after task τ_0. Right (first row): in subsequent tasks $\tau_1 \rightarrow \tau_3$, classical rehearsal approaches can access a decreasing amount of examples from their replay buffer: hence, overfitting shows as erosion of the initial boundaries. When applying Lipschitz-based constraints on replayed data (second row), small output variations are required around replay data, thus favoring less curved boundaries.

Mem./Speed/FAA trade-offs – [QKp5 - Applicability to real world scenario]

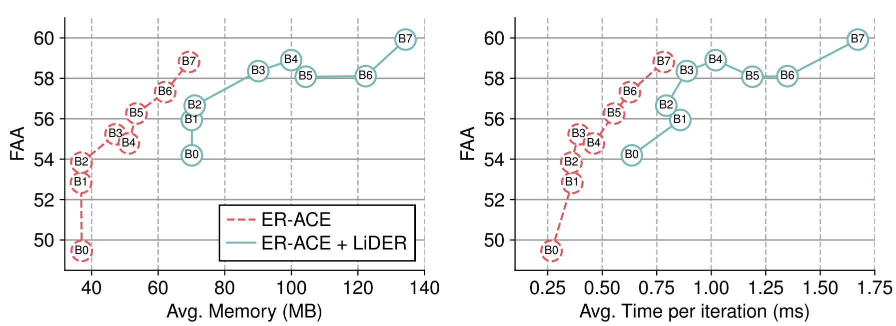

Figure B: *Memory footprint/accuracy* (left) and *speed/accuracy* (right) trade-offs of our approach in combination with ER-ACE. Results are reported for increasingly complex backbone networks (ranging from EffecientNET-B0 to -B7).

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

Sensitivity of α and β – [tK25 - Robustness to hyperparameters α and β]

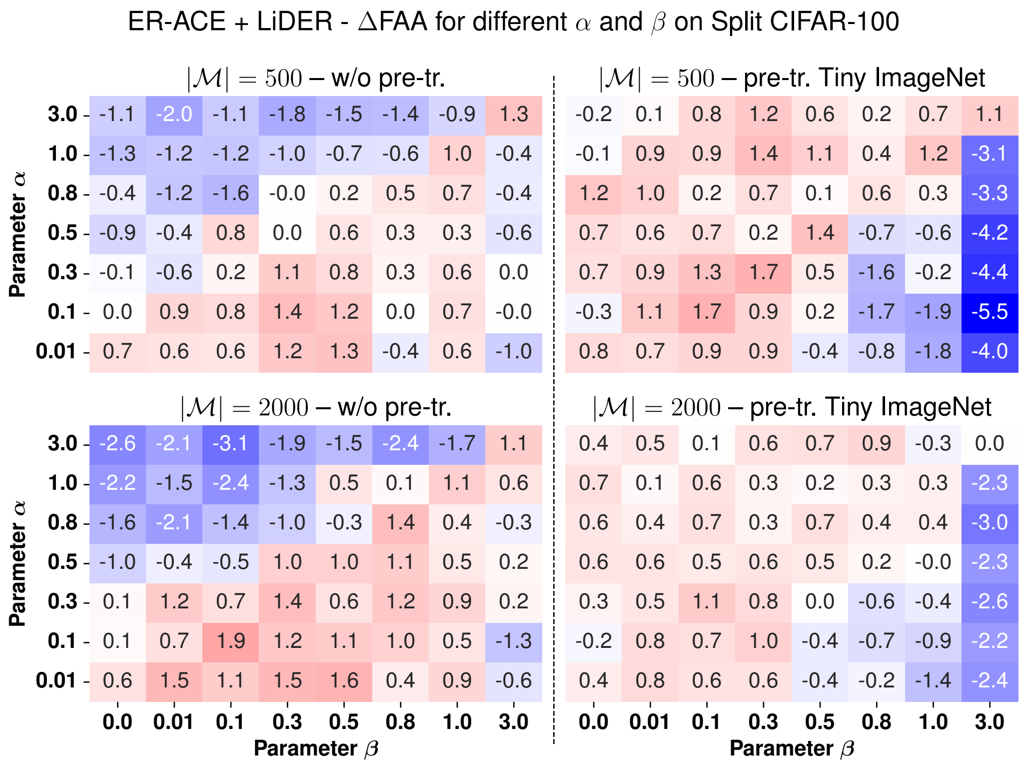

Figure C: Sensitivity analysis of ER-ACE + LiDER to the hyperparameters α and β on Split CIFAR-100. Results for different sizes of the memory buffer, with and without pre-training.