An analysis of best-practice strategies for replay and rehearsal in continual learning

Supplementary Material

6. Rationale

6.1. Additional Metrics: Forgetting

(a) Warm-start: Absolute forgetting.

(b) Warm-start: Normalized Forgetting.

(c) Generator reset: Absolute forgetting.

(d) Generator reset: Normalized Forgetting.(generator reset).

(e) Difference in absolute forgetting between (c.) and (a.).

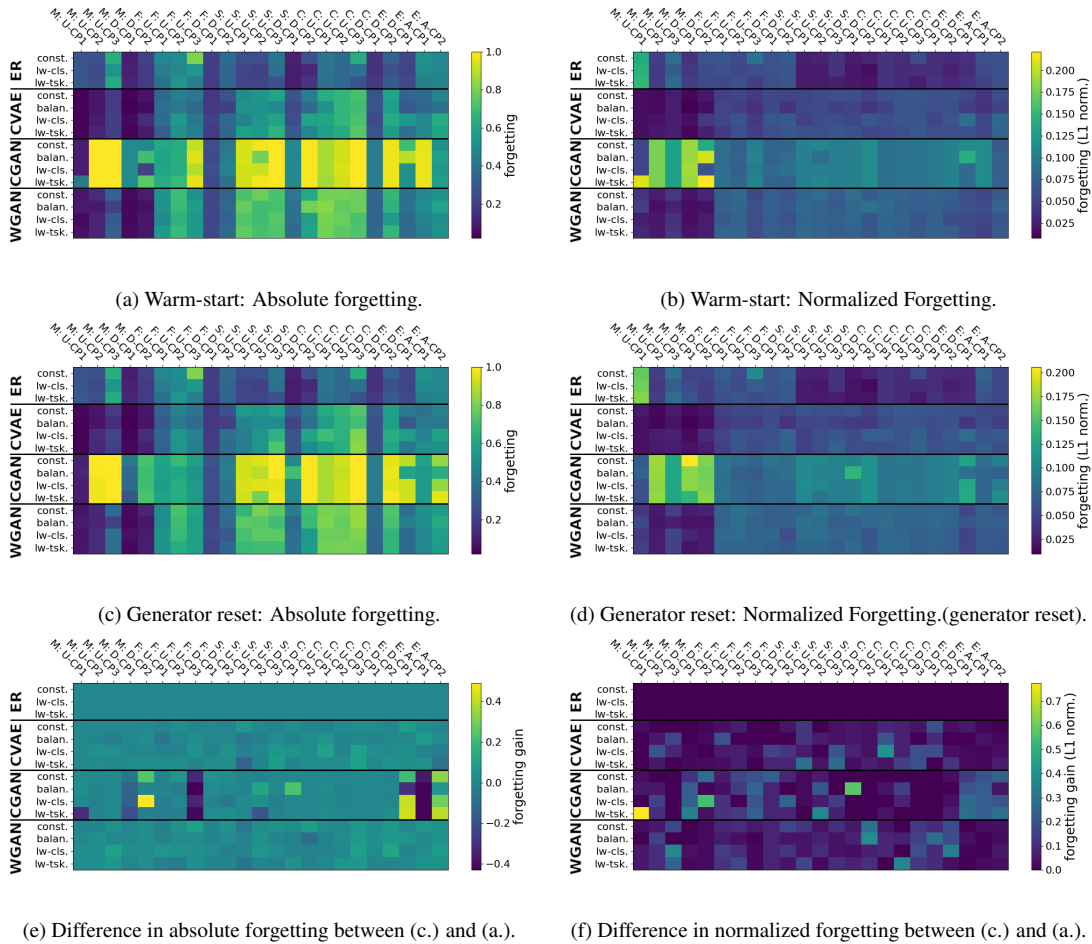(f) Difference in normalized forgetting between (c.) and (a.).

Figure 4. Averaged forgetting $F_{T_{\text{end}}}$ for all investigated datasets/splits. Each column represents a distinct task split on each dataset, whereas the first letter ("M", "F", "S", "C" stands for MNIST, Fashion-MNIST, SVHN and CIFAR-10), followed by the task split descriptor from Tab. 1. The deployed CL methods (4 groups * N rows) were modified as explained in Sec. 2.3: *const.* = constant, *balan.* = *balanced* (DGR exclusive), *lw-cls.* = loss-weighting by class count, *lw-tsk.* = loss-weighting by task count. The results for ER (top-most three rows) are re-used and serve as a baseline for the results in (c.-f.). Fig. (a.) and (b.) show the absolute and normalized (over each column vector) values for a warm-start, while Figs. (c.) and (d.) display the results when generators are reset after each task. Figs. (e.) and (f.) show the differences between resetting and re-using generators.