Low-Rank Continual Personalization of Diffusion Models

Łukasz Staniszewski*[1] Katarzyna Zaleska*[1] Kamil Deja [1,2]

[1]Warsaw University of Technology [2]IDEAS NCBR

*Equal Contribution

#TLDR

- We evaluate three different approaches to **initializing** and **merging** the individual adapters in **Continual Object and Style Personalization of Diffusion Models**.
- We show that **Naïve Continual Training** of LoRA leads to **catastrophic forgetting**, while other techniques can mitigate this issue, which originates from **weights' conflicts** in adapters that are sequentially trained.

Motivation

Continuous personalization of a Generative Model over several tasks leads to the **catastrophic forgetting** of previously encoded knowledge.

Recent approaches mitigate this issue by **merging** the adapters after all tasks, but this is **impractical** as tasks increase (about 8 LoRA weight matrices are equal to the size of all adapted parameters in SDXL).

We propose to study the effectiveness of merging techniques under the strict continual learning regime where the model with, at most, a single adapter is passed between tasks.

Methods

We evaluate a **Naïve continual fine-tuning** approach, where low-rank weights are fine-tuned from the previous task, with:

(1) Merge & Initialization. New task LoRA is initialized in a standard manner ($A \sim N(0, I), B = 0$). At the end of the task, the adapter is merged into the base model.

(2) Merge & Orthogonal Initialization. We initialize A_t weights as orthogonal to $A_{1..t-1}$ using Singular Value Decomposition (SVD). We decompose the i-th column as:

$$A^{(i)}_{1..t-1} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$$

and take the last row of $\mathbf{V}$, linked to the lowest singular value.

(3) Magnitude-based selection of LoRA weights. We adapt the MagMax method and select weights with the highest magnitude when comparing already merged adapters with the current one.

Analysis

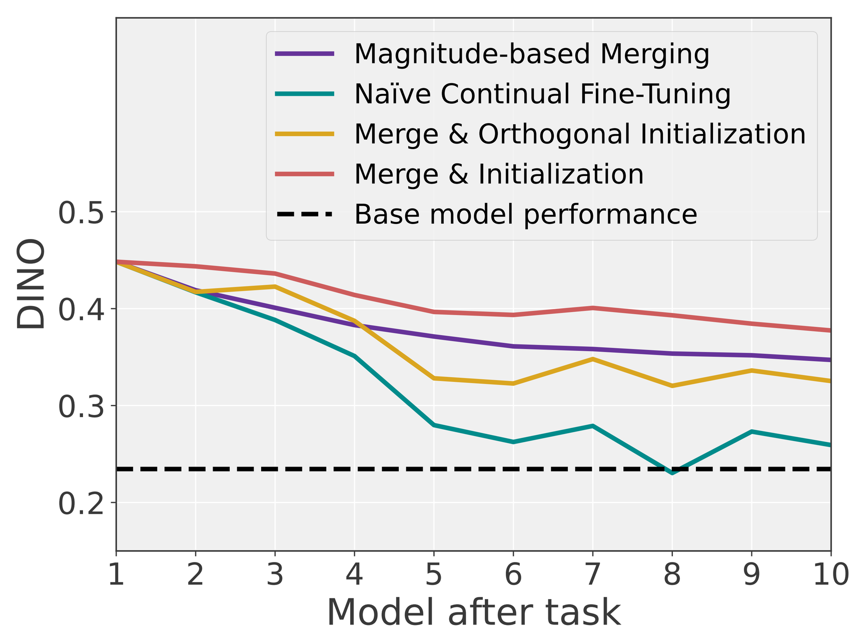


Figure 1. DINO score on the first task over continual fine-tuning on the next object tasks.

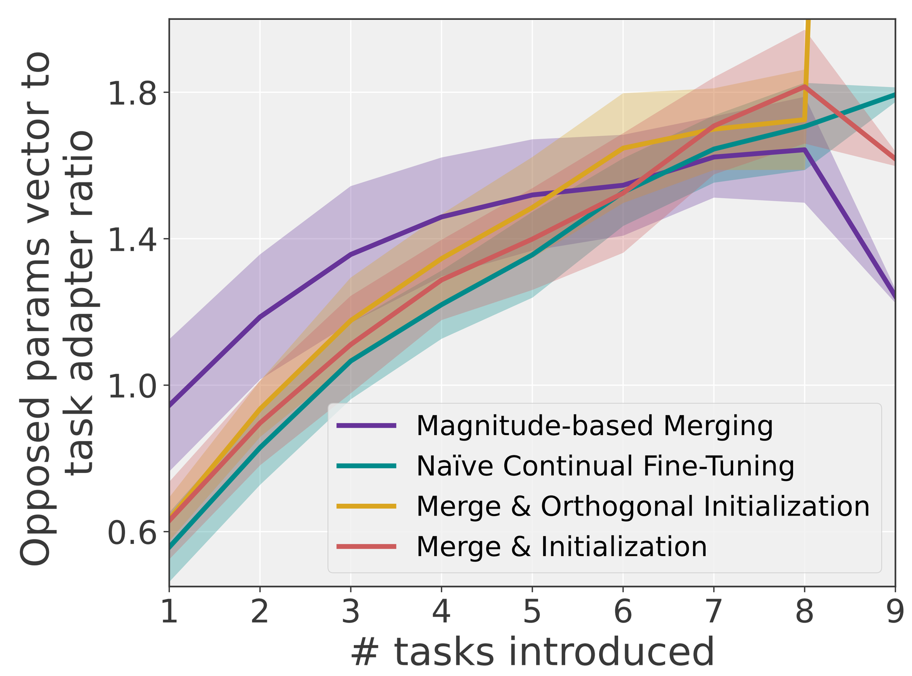
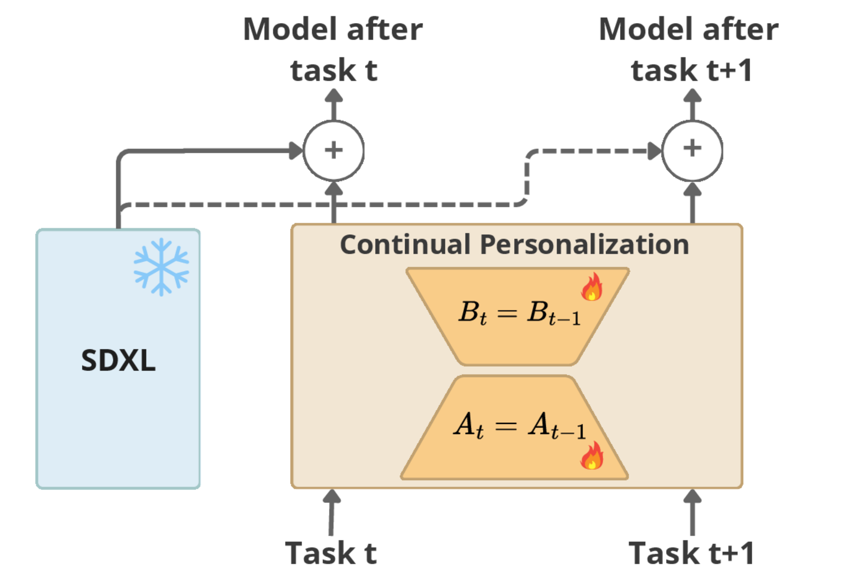


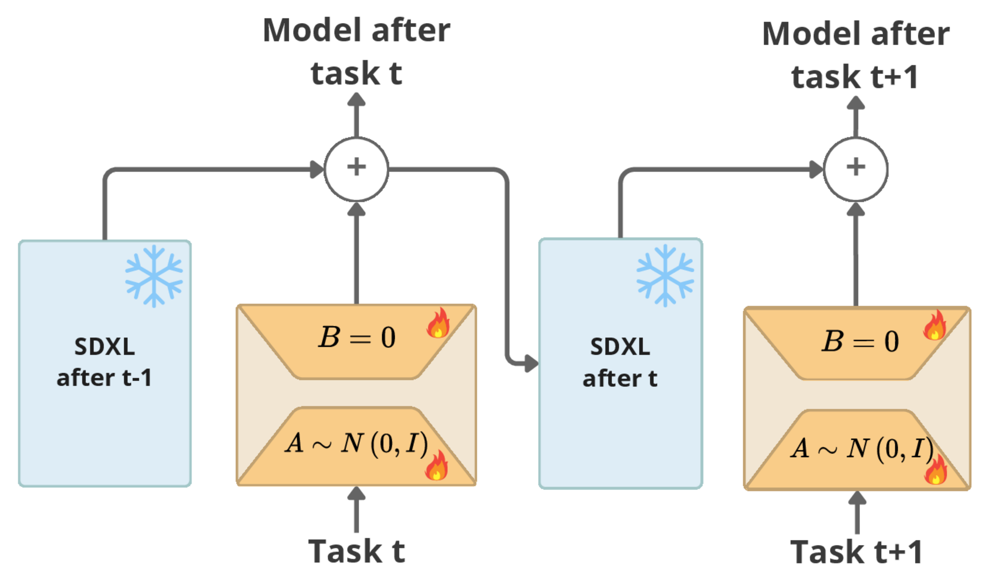
Figure 2. The ratio of opposed parameters vector norm to the task adapter norm.

Continual Diffusion Model Personalization

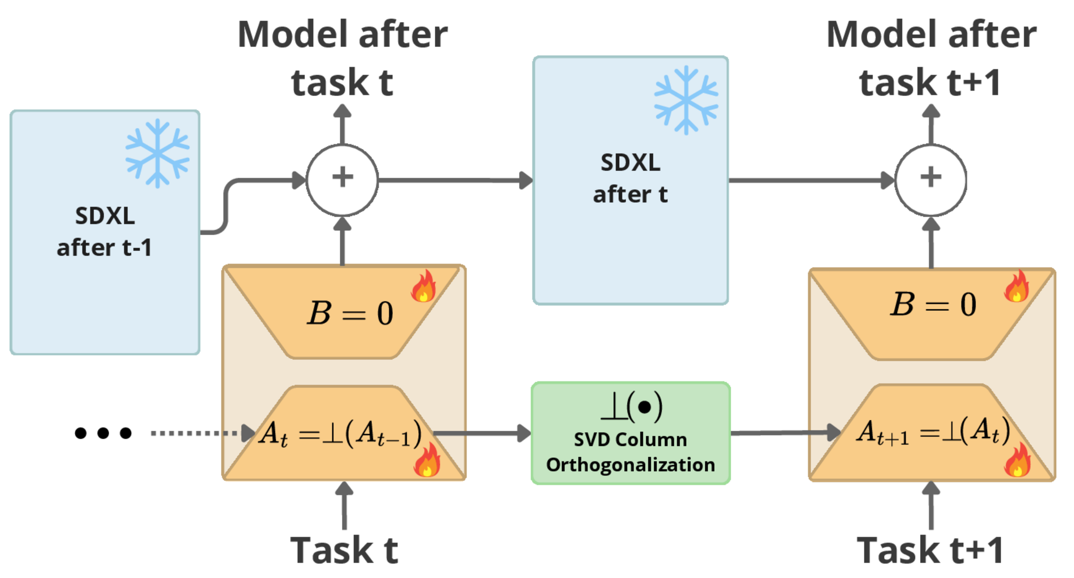
(Baseline) Naïve continual fine-tuning



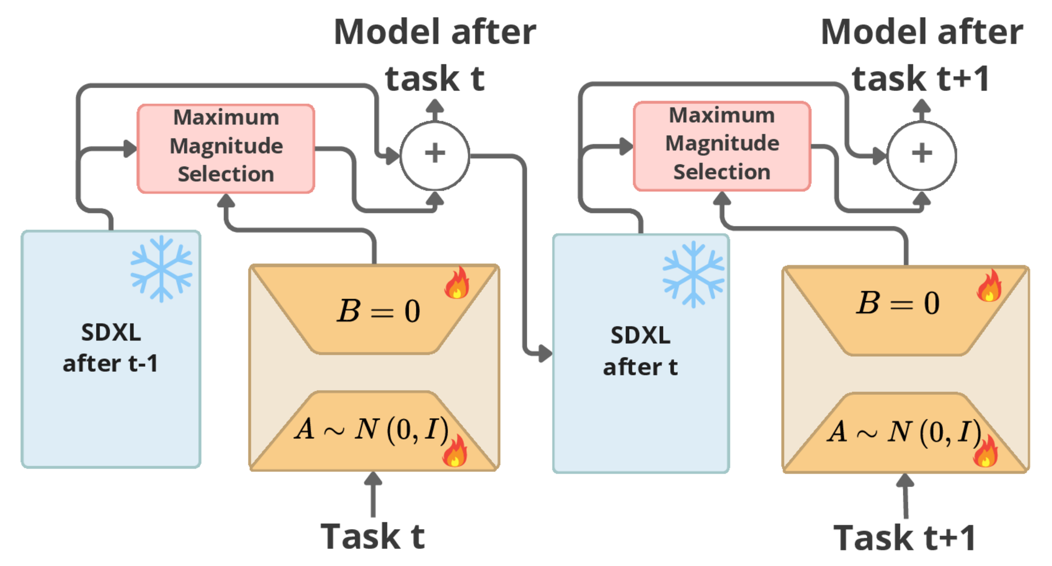
(1) Merge & Initialization



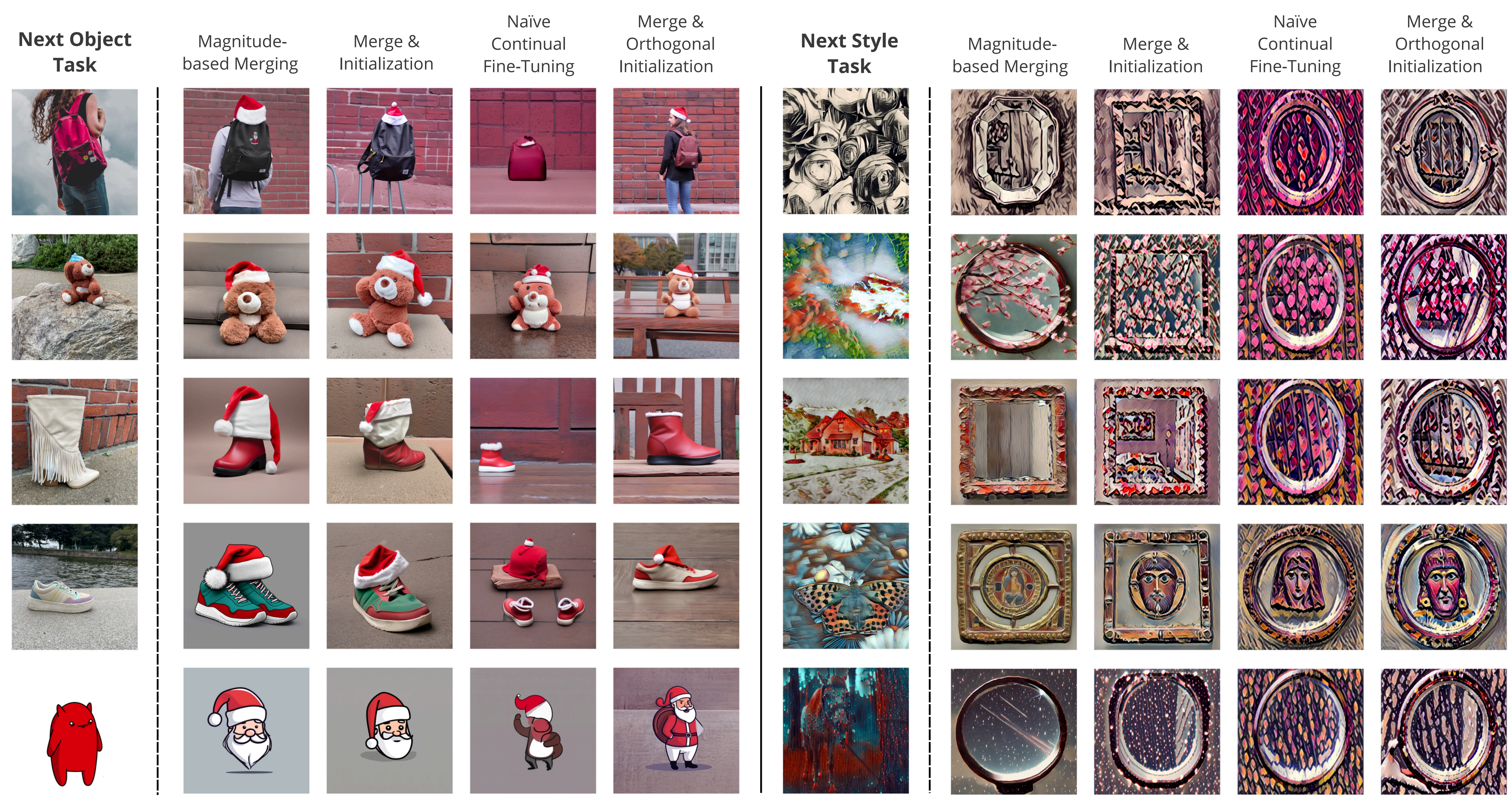
(2) Merge & Orthogonal Initialization



(3) Magnitude-based selection of LoRA weights



Examples



Results

Adapter fine-tuning method	Average Score $\overline{S}_T$ (↑) CLIP-I	Average Score $\overline{S}_T$ (↑) DINO	Average Forgetting $\overline{F}_T$ (↓) CLIP-I	Average Forgetting $\overline{F}_T$ (↓) DINO
Base model (reference)	0.586	0.304	-	-
Naïve Continual Fine-Tuning	$0.670_{\pm.013}$	$0.402_{\pm.029}$	$0.063_{\pm.013}$	$0.144_{\pm.039}$
Merge & Initialization	$\mathbf{0.675}_{\pm.005}$	$\mathbf{0.457}_{\pm.011}$	$\underline{0.026}_{\pm.003}$	$\underline{0.056}_{\pm.009}$
Merge & Orthogonal Initialization	$\underline{0.673}_{\pm.014}$	$0.403_{\pm.025}$	$0.072_{\pm.015}$	$0.162_{\pm.033}$
Magnitude-based Merging	$0.643_{\pm.002}$	$\underline{0.408}_{\pm.006}$	$\mathbf{0.018}_{\pm.002}$	$\mathbf{0.036}_{\pm.006}$

Table 1. Average Score and Average Forgetting for **Continual Object Personalization**.

Adapter fine-tuning method	Average Score $\overline{S}_T$ (↑) CSD	Average Score $\overline{S}_T$ (↑) DINO	Average Forgetting $\overline{F}_T$ (↓) CSD	Average Forgetting $\overline{F}_T$ (↓) DINO
Base model (reference)	0.088	0.146	-	-
Naïve Continual Fine-Tuning	$0.345_{\pm.032}$	$0.249_{\pm.010}$	$0.184_{\pm.039}$	$0.136_{\pm.013}$
Merge & Initialization	$\mathbf{0.385}_{\pm.019}$	$\mathbf{0.285}_{\pm.019}$	$\underline{0.131}_{\pm.019}$	$\underline{0.085}_{\pm.021}$
Merge & Orthogonal Initialization	$\underline{0.349}_{\pm.014}$	$\underline{0.252}_{\pm.012}$	$0.204_{\pm.013}$	$0.141_{\pm.013}$
Magnitude-based Merging	$0.289_{\pm.015}$	$0.240_{\pm.009}$	$\mathbf{0.093}_{\pm.013}$	$\mathbf{0.052}_{\pm.015}$

Table 2. Average Score and Average Forgetting for **Continual Style Personalization**.

Take-Away Points

- Adding multiple adapters in a Naïve way leads to the model which converges towards its base form, while all the evaluated techniques mitigate this issue.
- A high extent of the mutual interference between adapters during training is the origin of adapters' degradation.
- All presented approaches outperform the naïve approach in continual objects and styles personalization.

If you liked this...

See the full paper for more fascinating insights!

katarzyna.zaleska2.stud@pw.edu.pl

Supported by PL-Grid Infrastructure grant no. PLG/2024/017176.

lukasz.staniszewski.stud@pw.edu.pl