LoRA-X: Bridging Foundation Models with Training-Free Cross-Model Adaptation

Farzad Farhadzadeh, Senior Staff Engineer, Qualcomm Technologies Inc.

Debasmit Das, Staff Engineer, Qualcomm Technologies Inc.

Shubhankar Borse, Staff Engineer, Qualcomm Technologies Inc.

Fatih Porikli, Sr Director, Qualcomm Technologies Inc.

Snapdragon and Qualcomm branded products are products of Qualcomm Technologies, Inc. and/or its subsidiaries.

Confidential - Qualcomm Technologies, Inc. and/or its affiliated companies - May Contain Trade Secrets

Setup

Introduction:

- LoRA refers to low-rank adaptations of large language/vision models, designed to address the challenges of fine-tuning large models.
- LoRA can be applied to the cross-attention layers within a Stable Diffusion model, allowing it to relate image representations with the prompts that describe them.

LoRA advantages:

- Training is much faster, Compute requirements are lower, Trained weights are much smaller

Motivation:

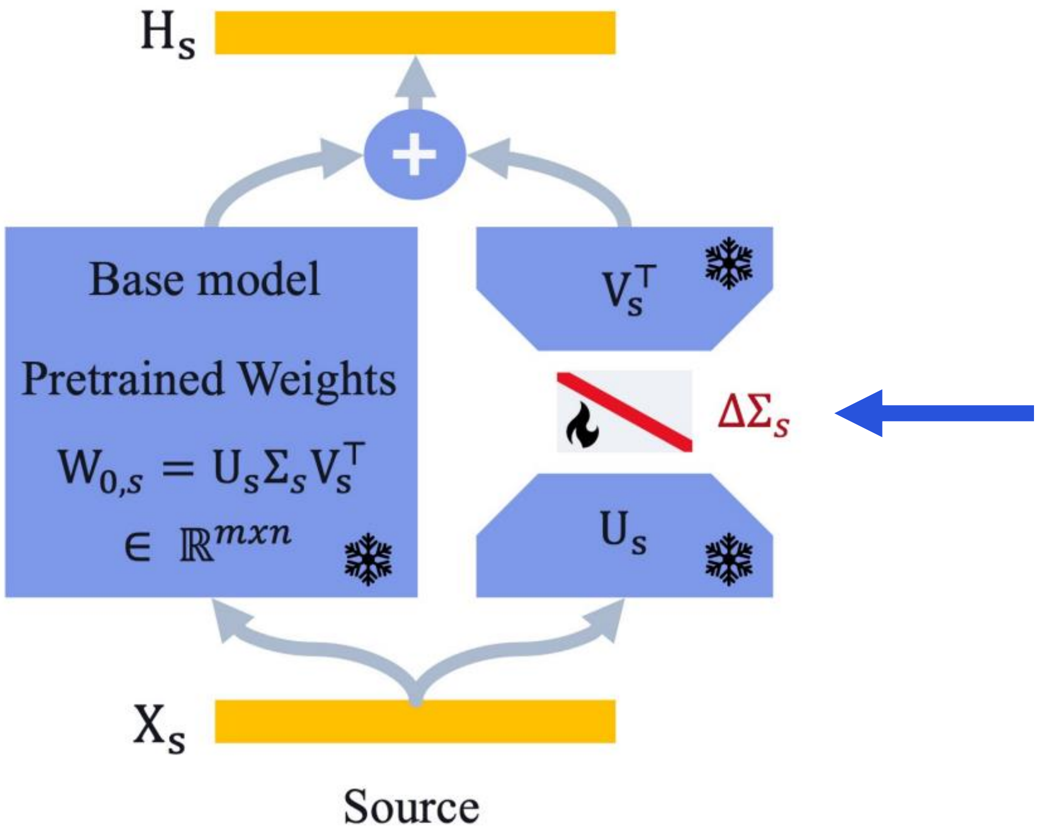
- Several LoRA adaptors are publicly available (without sharing dataset and training strategy)
- Stable Diffusion (SD) community keeps introducing new SD models

Goal:

- Distill knowledge from adapted Teacher model to Student model of different architecture and/or different sampling steps without having access to LoRA dataset.

Method

Train LoRA-X on source model



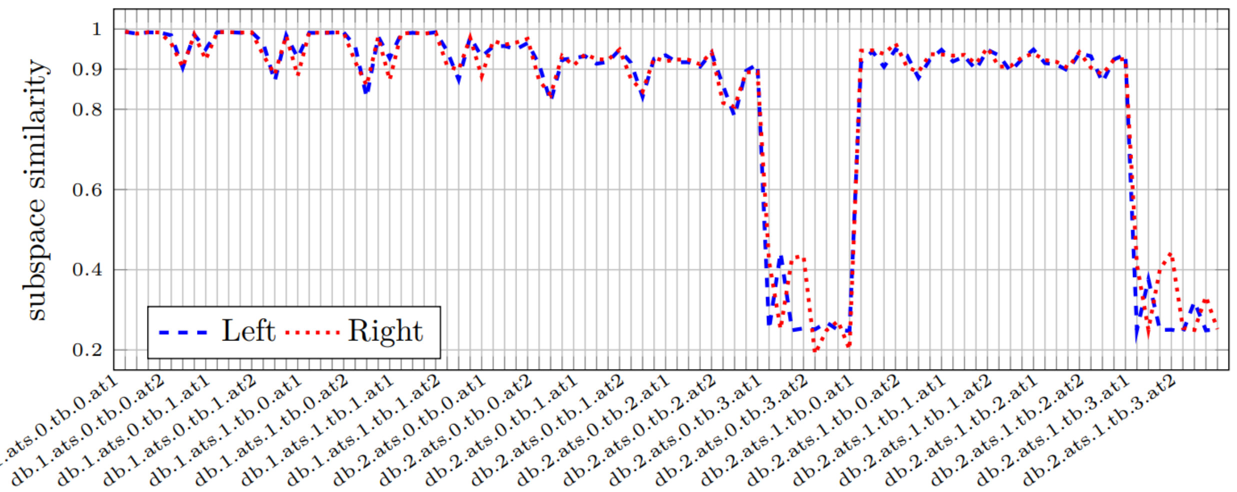
Only singular values are updated

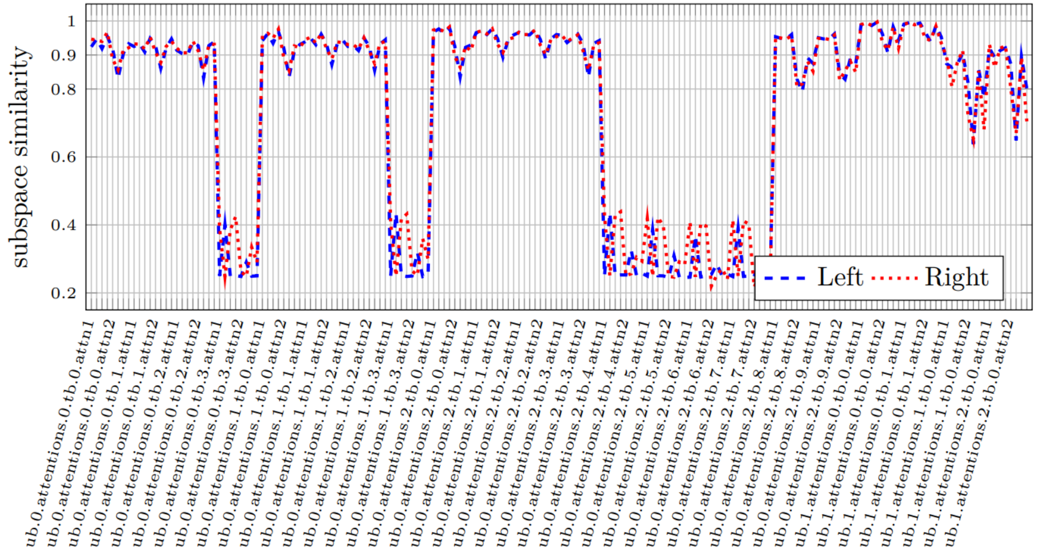
This is done to make sure adapter update lies in subspace of pretrained weights and ensure better transferability.

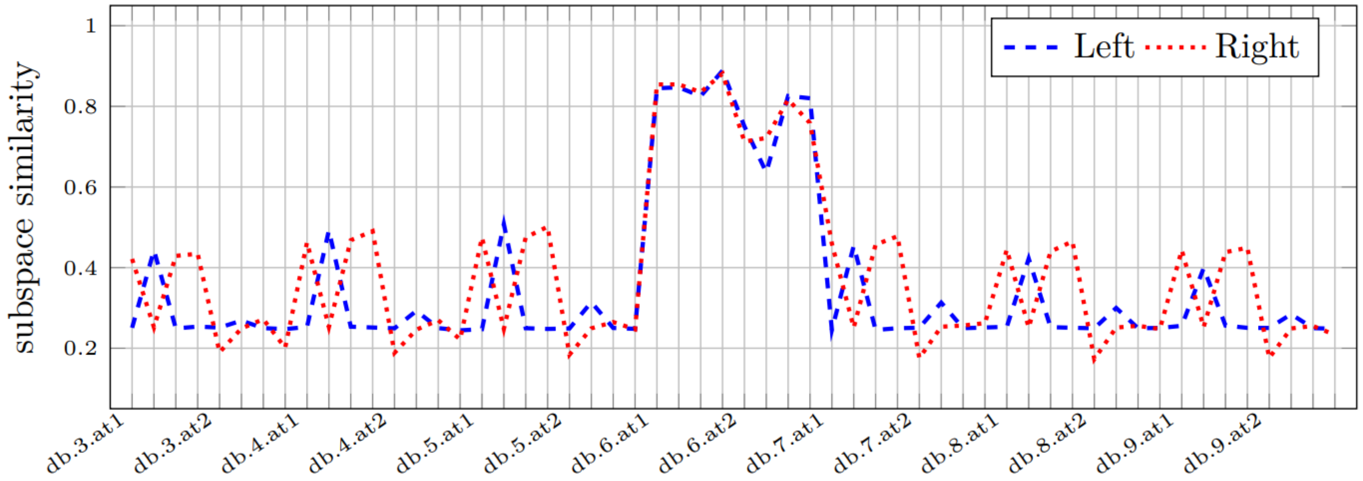
Method

Find subspace similarity between source & target modules

$$\Phi_l(A, B) = \frac{\|\boldsymbol{A}^\top \boldsymbol{B}\|_F^2}{\|\boldsymbol{A}^\top \boldsymbol{A}\|_F \|\boldsymbol{B}^\top \boldsymbol{B}\|_F} = \frac{\sum_i \sum_j \sigma_A^i \sigma_B^j \langle \boldsymbol{u}_A^i, \boldsymbol{u}_B^j \rangle^2}{\sqrt{\sum_i (\sigma_A^i)^2} \sqrt{\sum_i (\sigma_B^i)^2}}$$

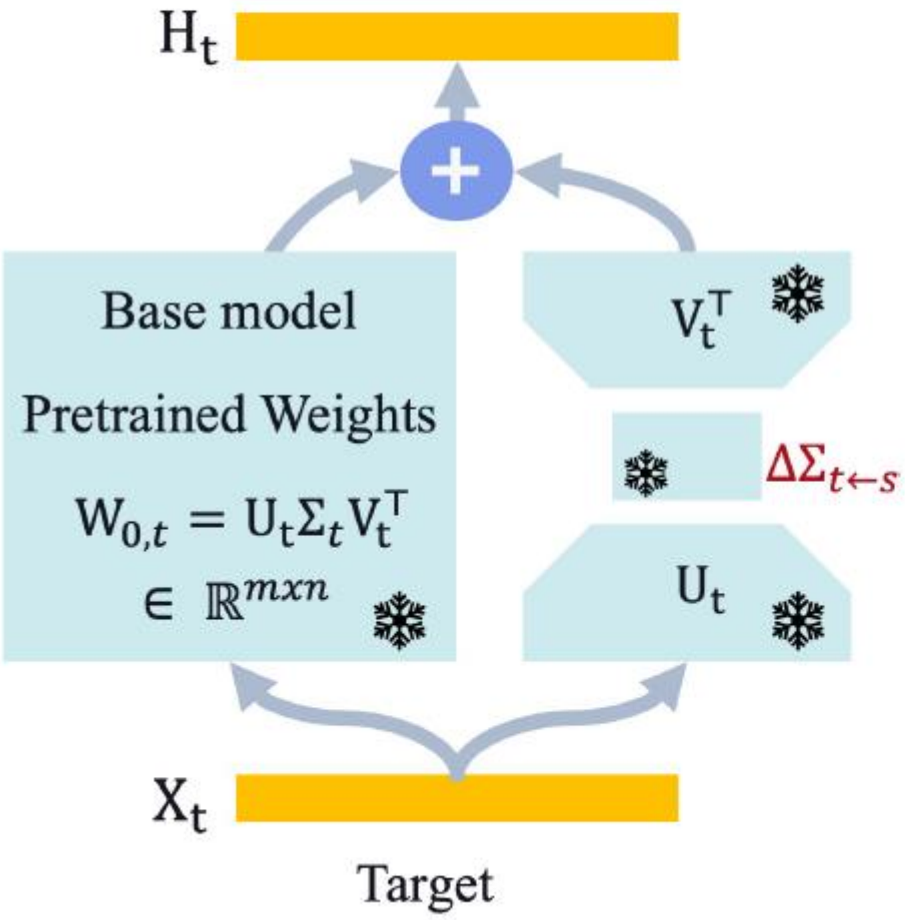






Method

Transfer LoRA-X from source and target



Given

$$\begin{aligned} \boldsymbol{W}_{s,0} &= \boldsymbol{U}_s \boldsymbol{\Sigma}_s \boldsymbol{V}_s^\top \\ \boldsymbol{W}_t &= \boldsymbol{U}_t \boldsymbol{\Sigma}_t \boldsymbol{V}_t^\top \\ \Delta\boldsymbol{W}_s &= \tilde{\boldsymbol{U}}_s \Delta\boldsymbol{\Sigma}_s \tilde{\boldsymbol{V}}_s^\top \end{aligned}$$

Same size LoRA across source & target

$$\Delta\boldsymbol{W}_{t\leftarrow s} = \boldsymbol{U}_t \boldsymbol{U}_t^\top \Delta\boldsymbol{W}_s \boldsymbol{V}_t \boldsymbol{V}_t^\top$$
$$= \boldsymbol{U}_t \boldsymbol{U}_t^\top \tilde{\boldsymbol{U}}_s \Delta\boldsymbol{\Sigma}_s \tilde{\boldsymbol{V}}_s^\top \boldsymbol{V}_t \boldsymbol{V}_t^\top = \boldsymbol{U}_t \Delta\boldsymbol{\Sigma}_{t\leftarrow s} \boldsymbol{V}_t^\top$$

$$m \neq m' \quad \tilde{\boldsymbol{U}}_s = \boldsymbol{U}_t \boldsymbol{U}_s^\top (\boldsymbol{U}_s \boldsymbol{U}_s^\top)^{-1} \boldsymbol{U}_s$$
$$n \neq n' \quad \tilde{\boldsymbol{V}}_s = \boldsymbol{V}_s (\boldsymbol{V}_s^\top \boldsymbol{V}_s)^{-1} \boldsymbol{V}_s^\top \boldsymbol{V}_t$$

Confidential - Qualcomm Technologies, Inc. and/or its affiliated companies - May Contain Trade Secrets

Qualitative Results

Adapter: "Painting", **Prompt:** 1) "astronaut floating in space" 2) "bear riding bike, traffic light" 3) "bird flying in the sky" 4) "elephant in a grassland" 5) "car on a winding road, mean headlights, thunderstorms, blue flames" 6) "horses eating grass, wooden hut".

Adapter: "Painting", **Prompt:** 1) "astronaut floating in space" 2) "bear riding bike, traffic light" 3) "bird flying in the sky" 4) "elephant in a grassland" 5) "horses eating grass, wooden hut" 6) "wild dolphins swimming".

Qualitative Results

Adapter: "Origami", Prompt: 1) "elephant" 2) "bird with spread wings" 3) "doberman dog" 4) "dragon" 5) "flower" 6) "truck".

Adapter: "Origami", Prompt: 1) "elephant" 2) "bird with spread wings" 3) "doberman dog" 4) "dragon" 5) "flower" 6) "truck".

Quantitative Results

Table 1: Evaluation of LoRA-X trained from scratch on base models versus training-free transferred LoRA-X from a source model into a target model. LoRA-X modifies the 320 largest singular values of the pre-trained weights. Results are averaged over 30 seeds.

Datasets	Base Model	Adapter	Training-Free	HPSv2 (↑)	LPIPS diversity (↑)	DINOv2 (↑)
BlueFire (900 images)	RealVis-v3.0	Trained		0.331	0.524	0.882
		Transferred	✓	0.332 (+**0.3**%)	0.540 (+**2.9**%)	
	SD Eff-v1.0	Trained		0.296	0.534	0.851
		Transferred	✓	0.307 (+**3.6**%)	0.538 (+**0.7**%)	
	RealVisXL-v3.0	Trained		0.319	0.484	0.947
		Transferred	✓	0.319 (**0.0**%)	0.456 (−**6.1**%)	
	SSD-1B	Trained		0.316	0.428	0.969
		Transferred	✓	0.300 (−**5.3**%)	0.392 (−**8.4**%)	
Paintings (630 images)	RealVis-v3.0	Trained		0.319	0.502	0.928
		Transferred	✓	0.329 (+**3.0**%)	0.441 (−**11.8**%)	
	SD Eff-v1.0	Trained		0.298	0.485	0.820
		Transferred	✓	0.292 (−**2.0**%)	0.476 (−**2.0**%)	
	RealVisXL-v3.0	Trained		0.333	0.467	0.945
		Transferred	✓	0.325 (−**2.5**%)	0.421 (−**9.6**%)	
	SSD-1B	Trained		0.319	0.409	0.961
		Transferred	✓	0.320 (+**0.3**%)	0.355 (−**13.2**%)	

Table 2: LoRA-X subspace constraint effect on transferability of style adapter. BlueFire dataset, SD-v1.5 as the source model and SD Eff-v1.0 as the target.

Method	Adapter	Rank	HPSv2 (↑)	LPIPS diversity (↑)	DINOv2 (↑)	Total size (MB)
LoRA-X	Trained	320	0.2958	0.5340	**0.8513**	0.16
	Transferred		0.3073 (+**3.7**%)	0.5376 (+**0.6**%)		
LoRA	Trained	32	0.3153	0.5049	0.8471	34.07
	Transferred		0.2466 (−**27.8**%)	0.4834 (−**4.4**%)		
	Trained	16	0.2652	0.5248	0.8266	17.08
	Transferred		0.2408 (−**10.1**%)	0.5224 (−**0.5**%)		
	Trained	1	0.2650	0.5312	0.8228	1.15
	Transferred		0.2355 (−**12.5**%)	0.5274 (−**0.7**%)		

Quantitative Results

Table 3: Transferability of style adapters DoRA & FouRA. For DoRA, SDXL is the source model and SSD-1B is the target model. For FouRA, SD-v1.5 is the source model and SD Eff-v1.0 is the target model.

Method	Adapter	Rank	Dataset	HPSv2 (↑)	LPIPS diversity (↑)	DINOv2 (↑)
DoRA	Trained	8	Paintings	0.3042	0.4624	**0.9138**
	Transferred			0.2764 (−**9.1**%)	0.4526 (−**2.1**%)	
DoRA	Trained	8	Origami	0.2491	0.3408	**0.9441**
	Transferred			0.2224 (−**10.7**%)	0.3073 (−**9.8**%)	
FouRA	Trained	64	Paintings	0.3034	0.4686	**0.9153**
	Transferred			0.2891 (−**4.7**%)	0.4446 (−**5.1**%)	

Table 4: Evaluation of training-free transferred LoRA-X from SSD-1B to SDXL versus LoRA-X trained on SDXL from scratch using BlueFire dataset using our training-free transfer method and training-based X-adapter. Wall clock time is measured on A100 GPU

Method	Adapter	HPSv2 (↑)	LPIPS diversity (↑)	DINOv2 (↑)	Time (↓)
LoRA-X	Trained	0.306	0.422	0.953	3.7s
	Transferred	0.279 (−**9.5**%)	0.433 (+**2.6**%)		
X-Adapter	Trained	0.306	0.422	0.892	17.1 s
	Transferred	0.282 (−**7.8**%)	0.406 (−**3.7**%)		

Confidential - Qualcomm Technologies, Inc. and/or its affiliated companies - May Contain Trade Secrets

Conclusion

Training free transfer of LoRA-X produces similar performance as training from scratch

LoRA-X transfer also works better across other adapter types

LoRA-X transfer also works better than existing transfer types like X-adapter, Copying

Confidential - Qualcomm Technologies, Inc. and/or its affiliated companies - May Contain Trade Secrets

Thank you

Nothing in these materials is an offer to sell any of the components or devices referenced herein.

© Qualcomm Technologies, Inc. and/or its affiliated companies. All Rights Reserved.

Qualcomm and Snapdragon are trademarks or registered trademarks of Qualcomm Incorporated.
Other products and brand names may be trademarks or registered trademarks of their respective owners.

References in this presentation to “Qualcomm” may mean Qualcomm Incorporated, Qualcomm Technologies, Inc., and/or other subsidiaries or business units within the Qualcomm corporate structure, as applicable. Qualcomm Incorporated includes our licensing business, QTL, and the vast majority of our patent portfolio. Qualcomm Technologies, Inc., a subsidiary of Qualcomm Incorporated, operates, along with its subsidiaries, substantially all of our engineering, research and development functions, and substantially all of our products and services businesses, including our QCT semiconductor business.

Snapdragon and Qualcomm branded products are products of Qualcomm Technologies, Inc. and/or its subsidiaries.
Qualcomm patents are licensed by Qualcomm Incorporated.

Follow us on:

For more information, visit us at qualcomm.com & qualcomm.com/blog