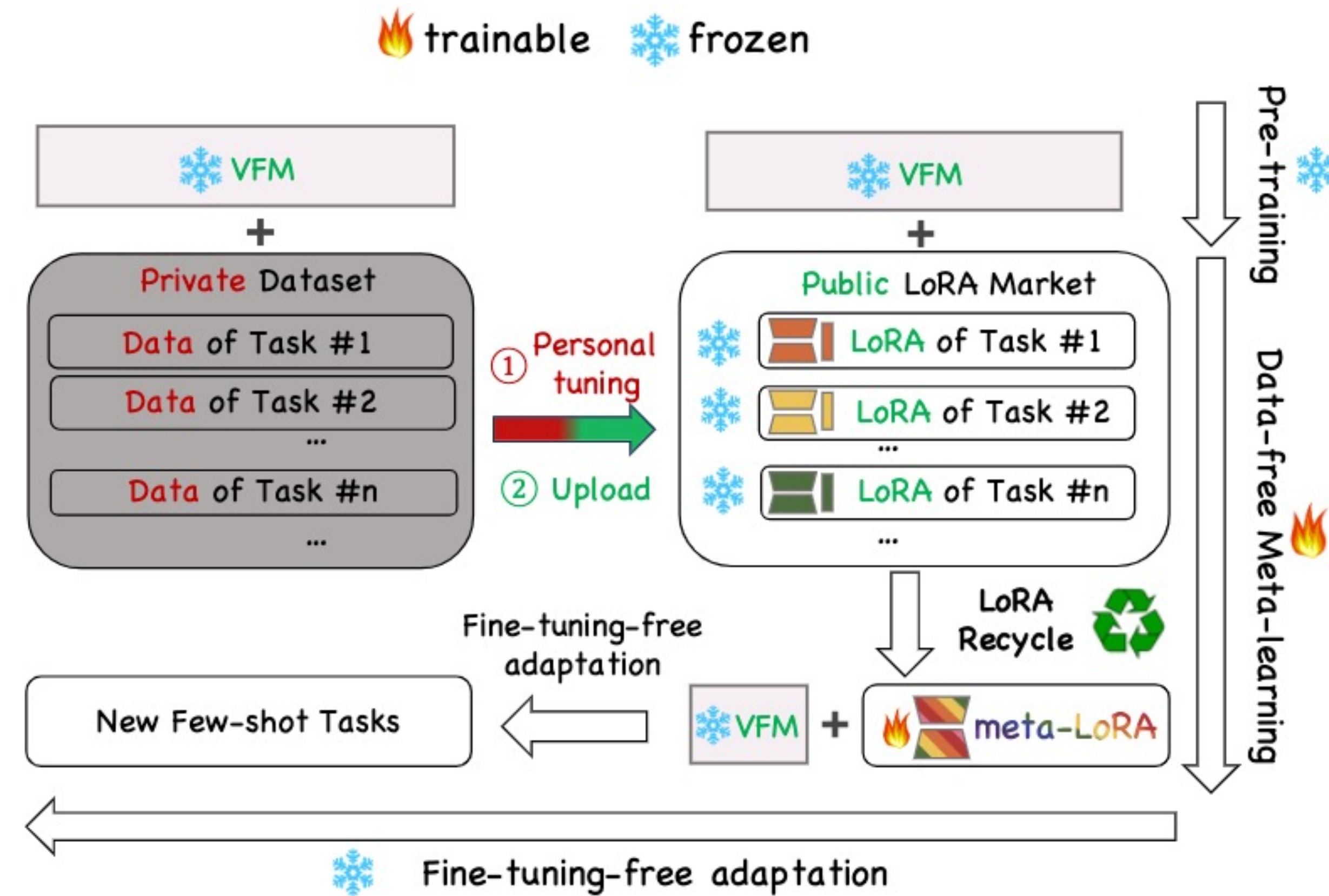


LoRA Recycle: Unlocking Tuning-Free Few-Shot Adaptability in Visual Foundation Models by Recycling Pre-Tuned LoRAs

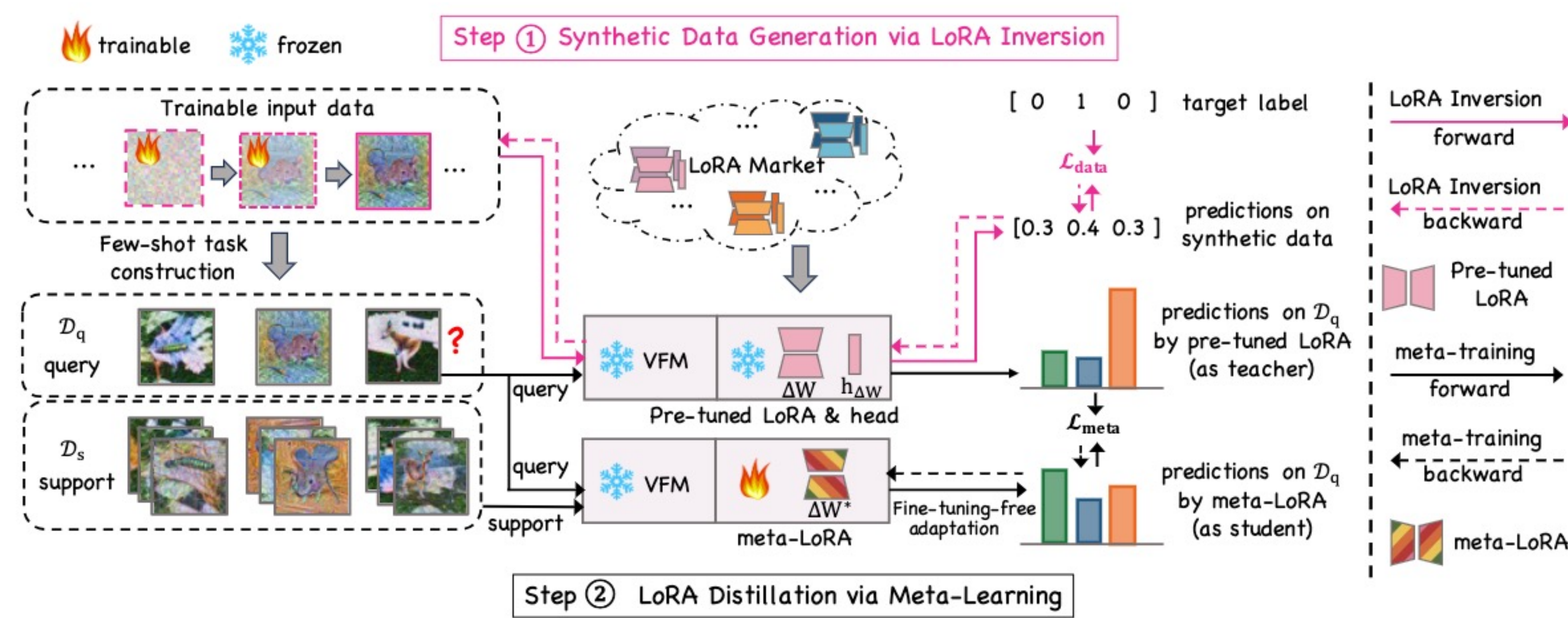
Zixuan Hu¹, Yongxian Wei², Li Shen³, Chun Yuan², Dacheng Tao¹



Motivation



LoRA Recycle

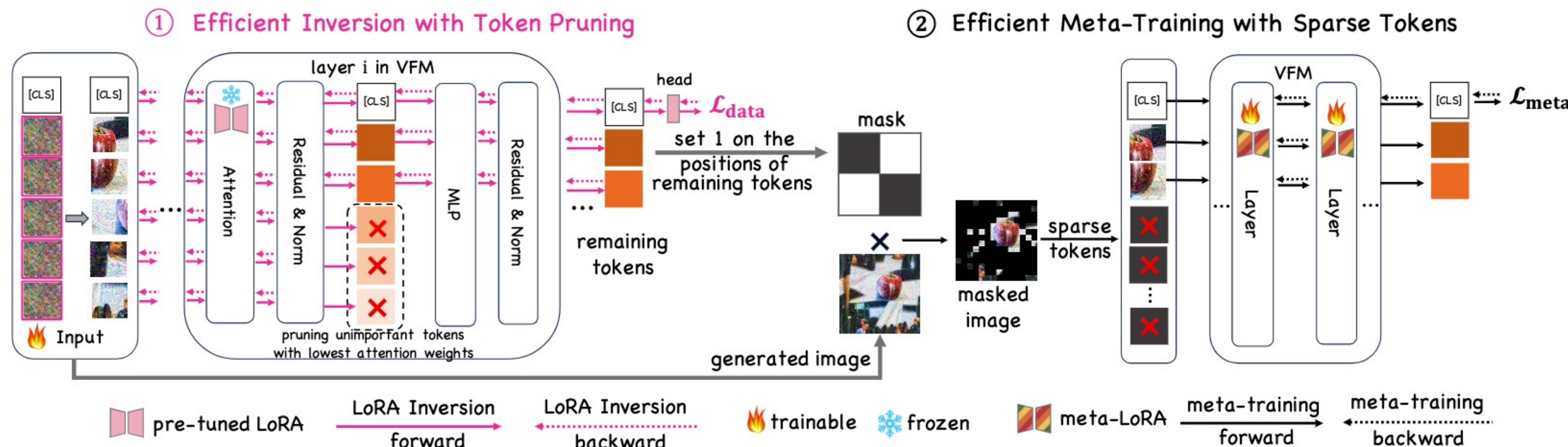


Experiments

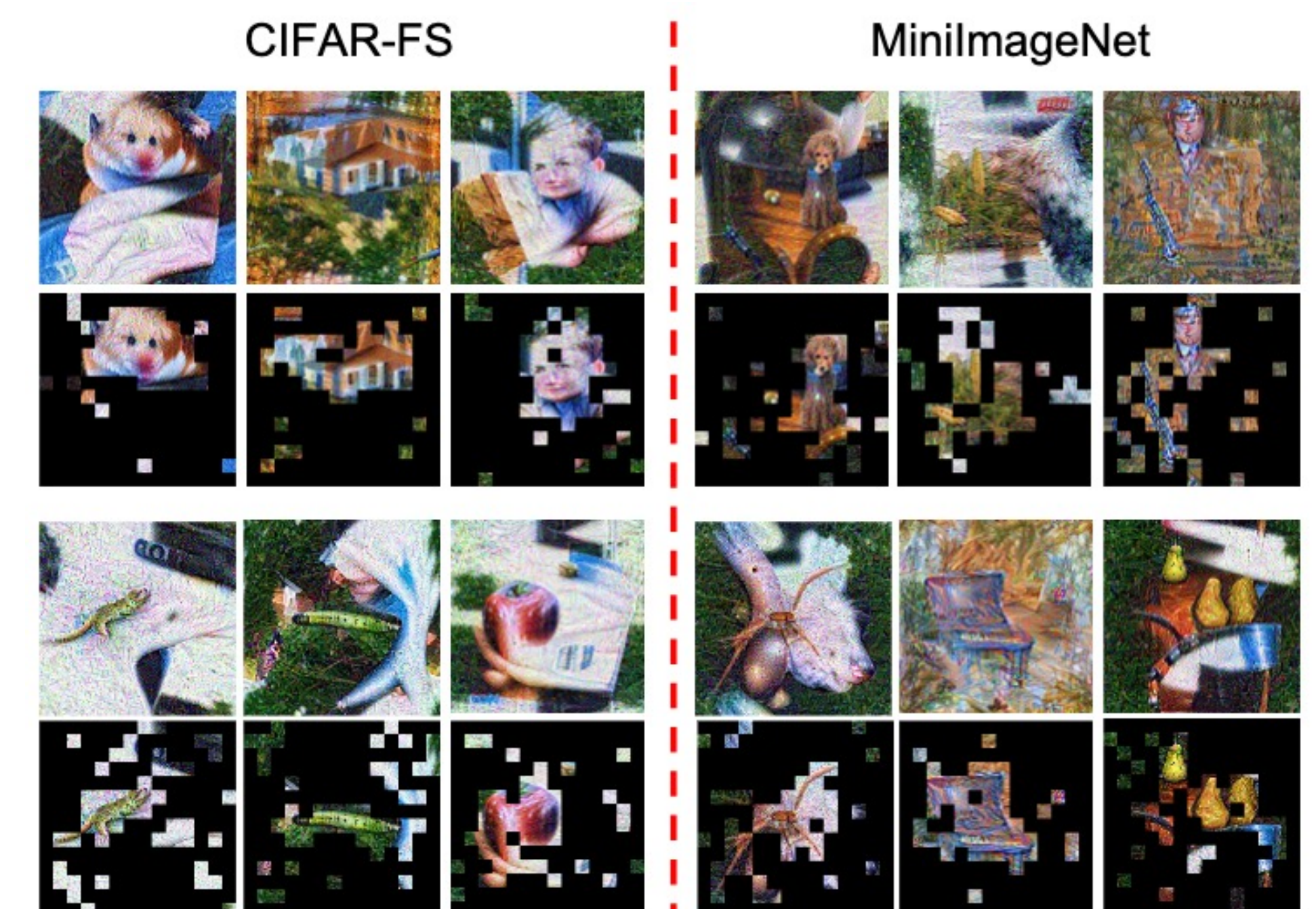
Few-Shot Adaptation

Method		CIFAR-FS		MiniImageNet		
		5-way 1-shot	5-way 5-shot	5-way 1-shot	5-way 5-shot	
FT	Full Finetuning	22.81	28.33	21.16	23.60	
	Linear-probe	80.06	95.49	82.04	94.12	
	LoRA + Linear	79.29	95.43	82.00	94.83	
	P > M > F	79.54	95.62	82.77	95.12	
	LoRAs Avg + Linear	80.25	96.07	83.59	95.43	
	MOLE	80.31	96.11	83.53	95.41	
	LoRAHub	81.23	96.24	83.68	95.72	
FTF	NN	78.06	94.09	81.08	93.85	
	LoRAs Avg + NN	79.37	93.45	81.72	94.64	
	CMAL	81.02	93.59	81.89	94.81	
	LoRA Recycle	89.69	97.05 ^{(+0.81%) (+2.96%)}	88.60 ^{(+4.92%) (+6.71%)}	96.12	
	LoRA Recycle ₂₅	91.03 ^{(+9.80%) (+10.01%)}	96.53	87.51	96.25 ^{(+0.53%) (+1.41%)}	
	LoRA Recycle ₅₀	90.91	96.08	87.21	95.85	
	LoRA Recycle ₇₅	89.70	96.69	87.36	96.05	
Token Pruning Strategy		5w 1s	5w 5s	Throughput (its/s) ↑	FLOPs (G) ↓	GPU Mem (GB) ↓
{0: 0.0}		89.69	97.05	5.56	50.59	8.74
{11: 0.75}		89.43	96.72	5.81 (+4%)	48.51 (-4%)	8.63 (-1%)
{8: 0.75}		82.27	95.69	6.22 (+12%)	39.14 (-23%)	8.07 (-8%)
{6: 0.75}		81.08	95.52	7.15 (+29%)	32.89 (-35%)	7.69 (-12%)
{3: 0.3, 6: 0.3, 8: 0.3, 11: 0.3}		84.17	96.12	6.13 (+10%)	40.00 (-21%)	8.08 (-8%)

Double-Efficient Mechanism



Visualization of Synthetic Data



- Thanks to the modularity of LoRA, users can upload locally tuned LoRAs to public repositories without exposing original training data.
- LoRA Recycle distills a meta-LoRA from these LoRAs without needing their original training data. The VFM, once equipped with the meta-LoRA, is empowered to solve new few-shot tasks in a single forward pass without further fine-tuning.