



# ICML

International Conference  
On Machine Learning

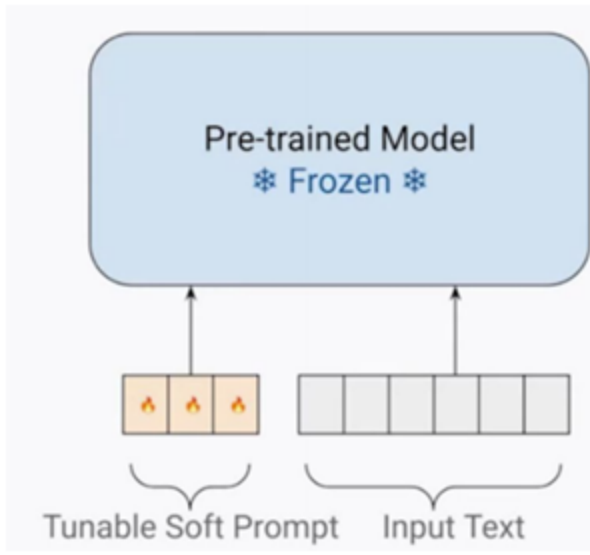


# Advancing Analytic Class-Incremental Learning through Vision-Language Calibration

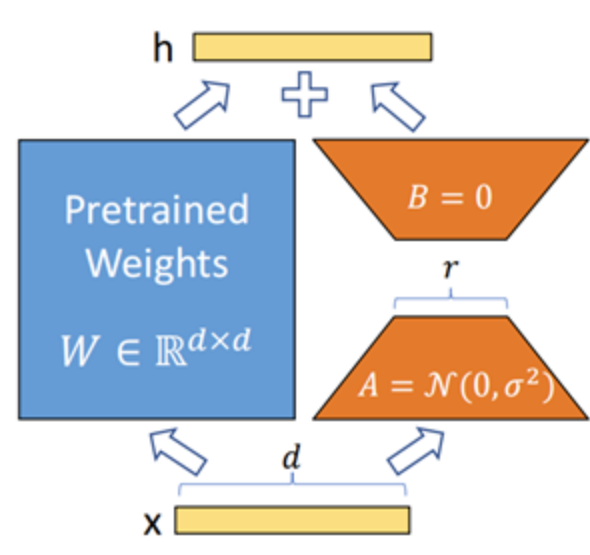
Binyu Zhao

Harbin Institute of Technology

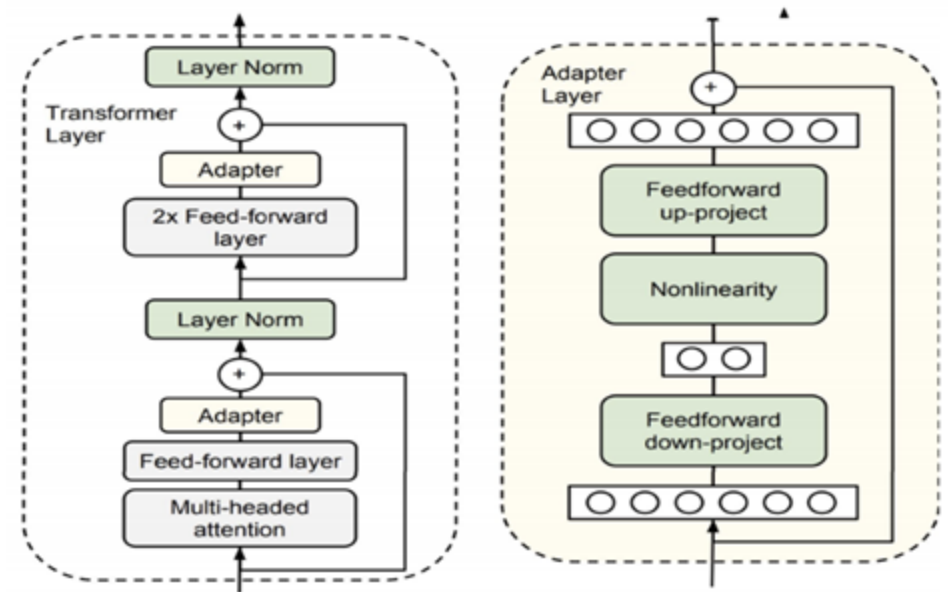
# Mainstream Iterative Tuning: High Cost & Progressive Degradation



Prompt

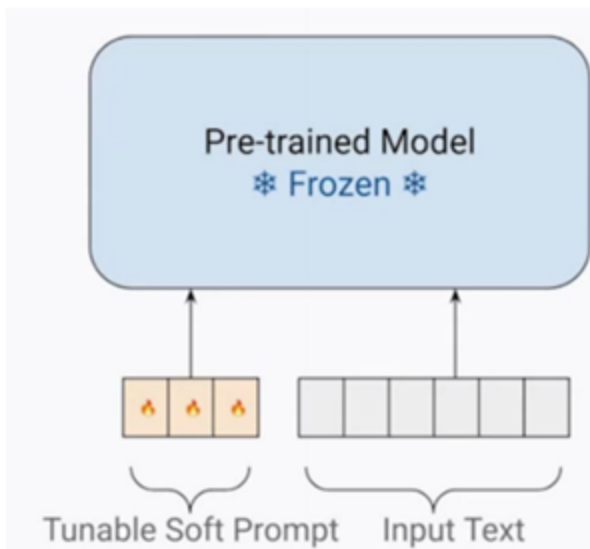


LoRA

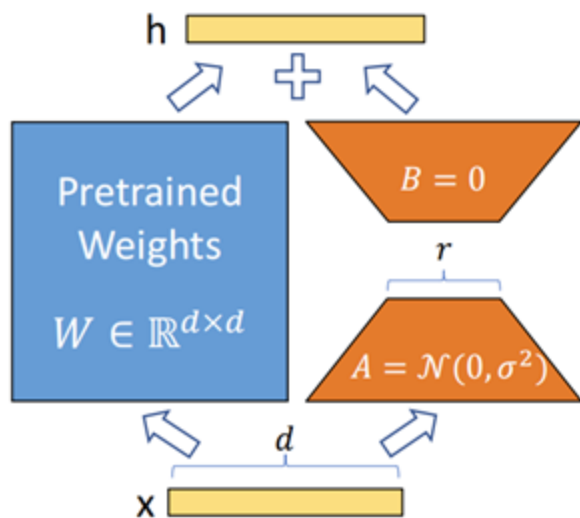


Adapter

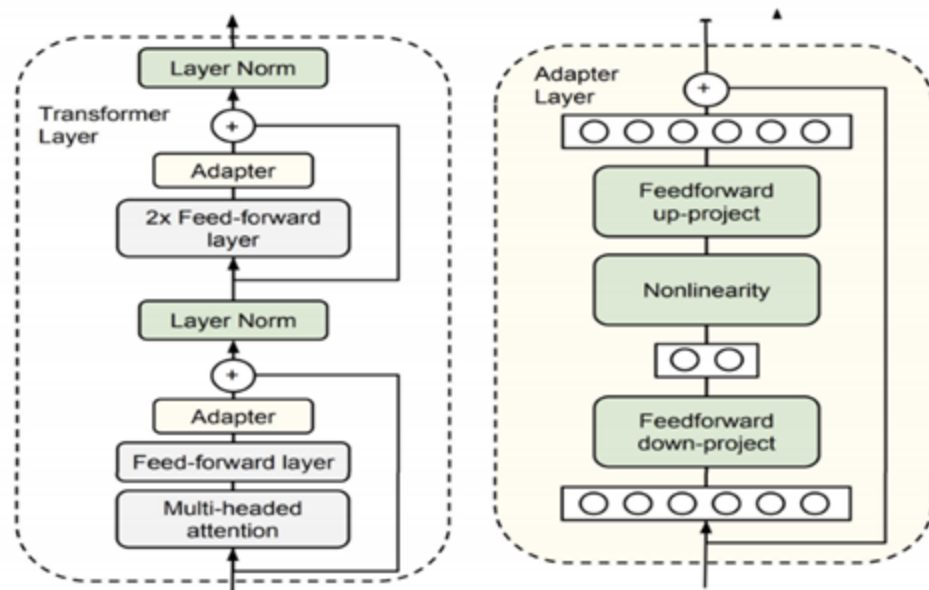
# Mainstream Iterative Tuning: High Cost & Progressive Degradation



Prompt

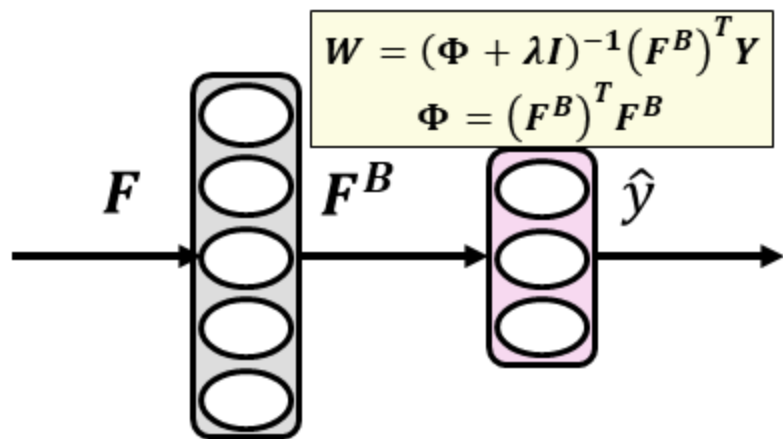


LoRA



Adapter

# Analytic Learning: Zero Forgetting & Closed-form Update



$$\min_W \| \mathbf{F}^B W - Y^{OH} \|_F^2 + \lambda \| W \|_F^2$$

ViT-Adapter Subspace  
(Representation Rigidity!)

Feature/Topology Gap

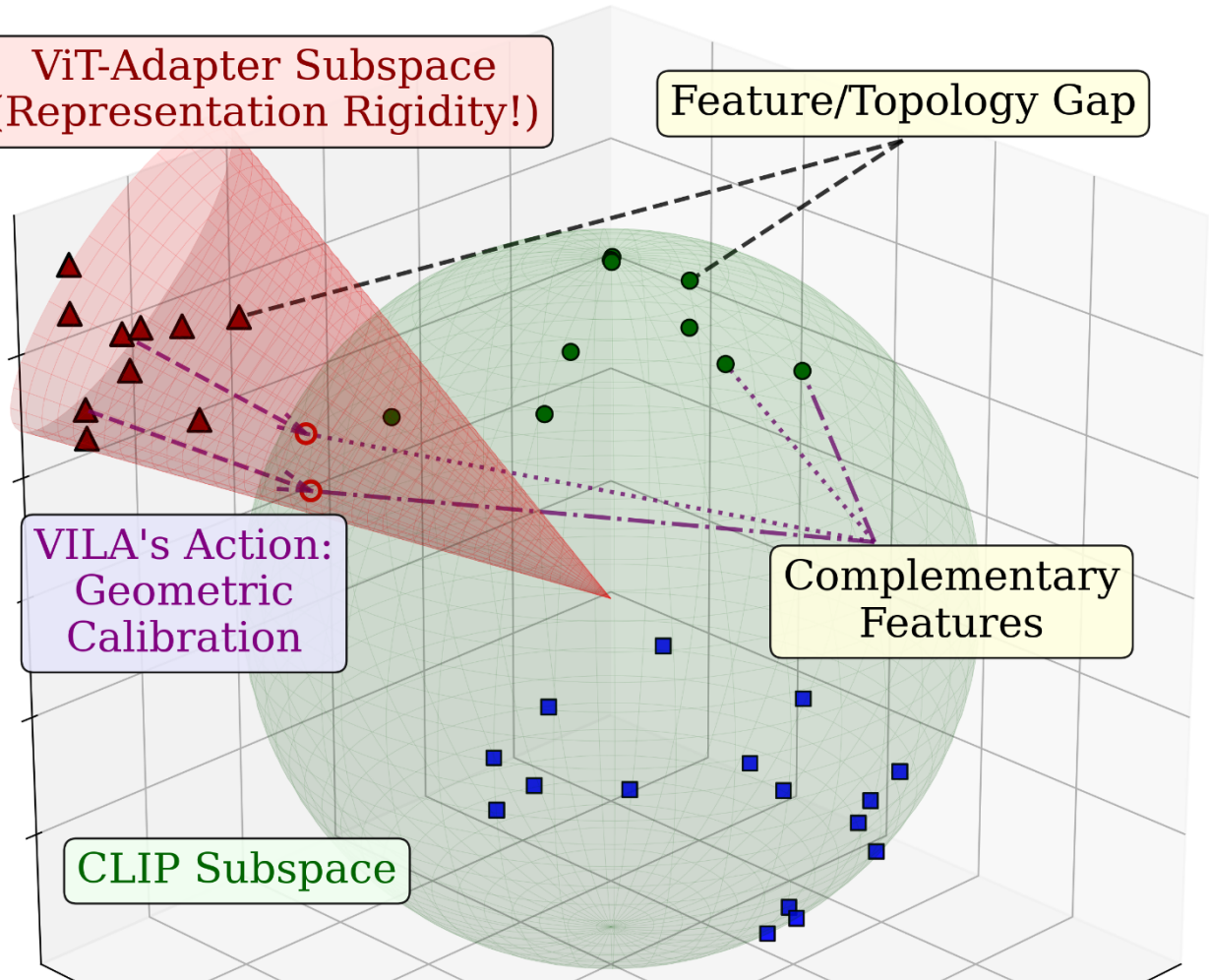
VILA's Action:  
Geometric  
Calibration

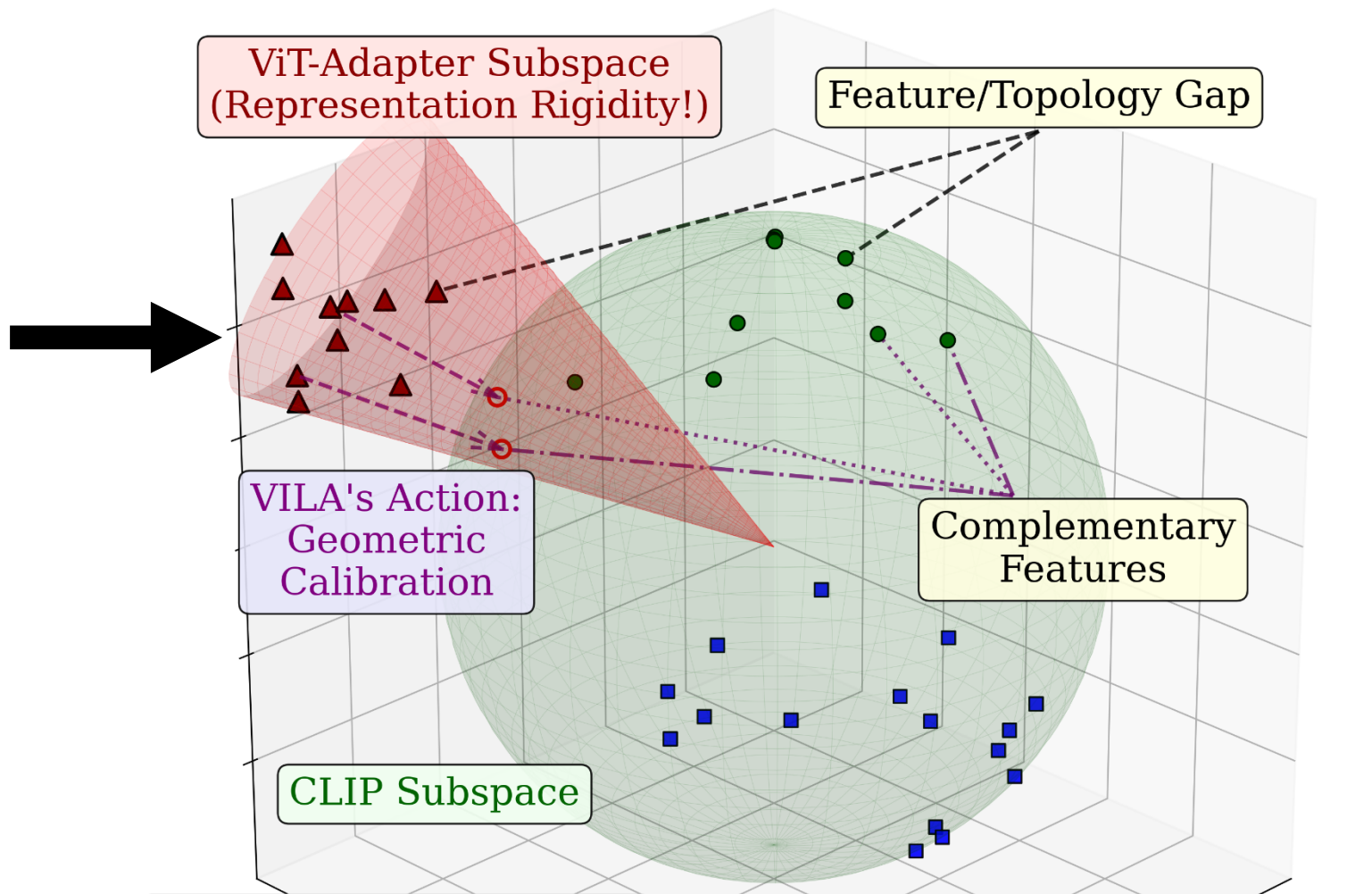
Complementary  
Features

CLIP Subspace

- ▲ Task 1 ViT-Adapter Features
- Task 1 Projected ViT-Adapter Features
- Task 1 CLIP Features
- Task 2 CLIP Features

Obs1 (Collapse in long sequences):  
The ViT-Adapter subspace is narrow  
Obs2 (Single branch conflict):  
Shape of ViT-Adapter/CLIP subspace  
Obs3 (CLIP geometric misalignment):  
Subspace gap & topological difference





ViT-Adapter Subspace  
(Representation Rigidity!)

Feature/Topology Gap

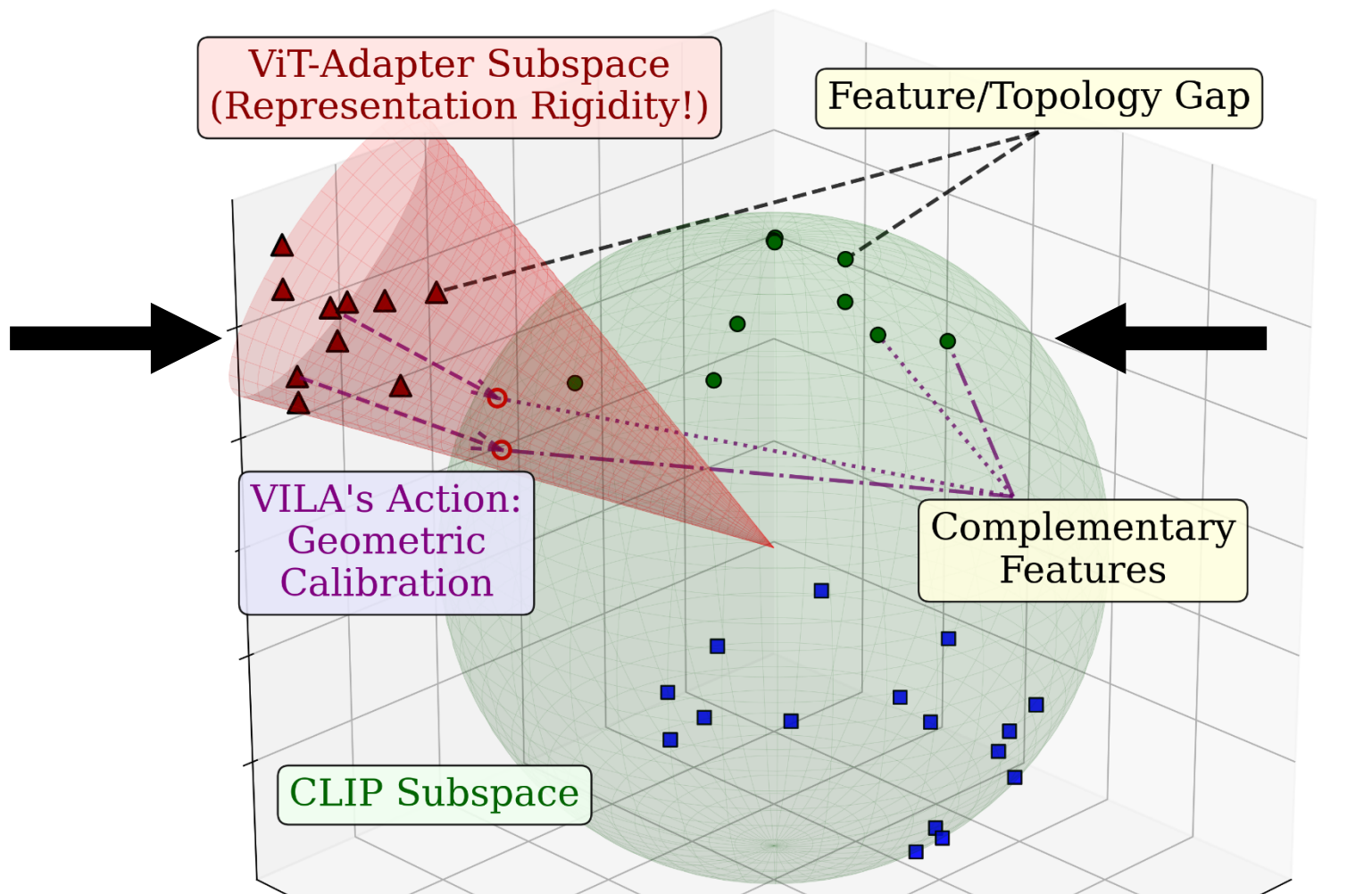
VILA's Action:  
Geometric  
Calibration

Complementary  
Features

CLIP Subspace

- ▲ Task 1 ViT-Adapter Features
- Task 1 Projected ViT-Adapter Features
- Task 1 CLIP Features
- Task 2 CLIP Features

Obs1 (Collapse in long sequences):  
The ViT-Adapter subspace is narrow  
Obs2 (Single branch conflict):  
Shape of ViT-Adapter/CLIP subspace  
Obs3 (CLIP geometric misalignment):  
Subspace gap & topological difference



ViT-Adapter Subspace  
(Representation Rigidity!)

Feature/Topology Gap

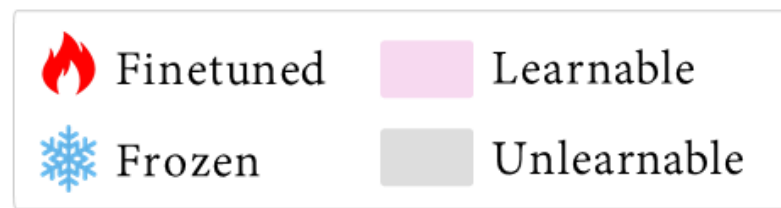
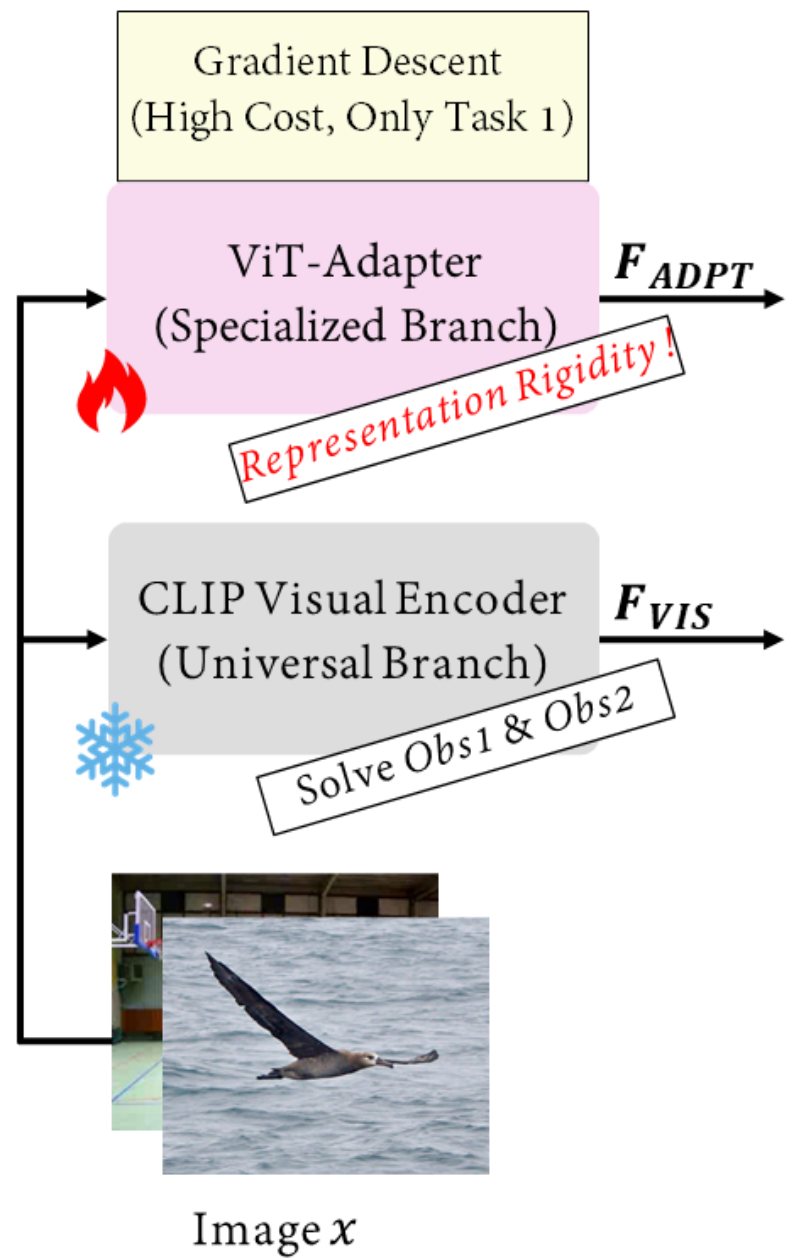
VILA's Action:  
Geometric  
Calibration

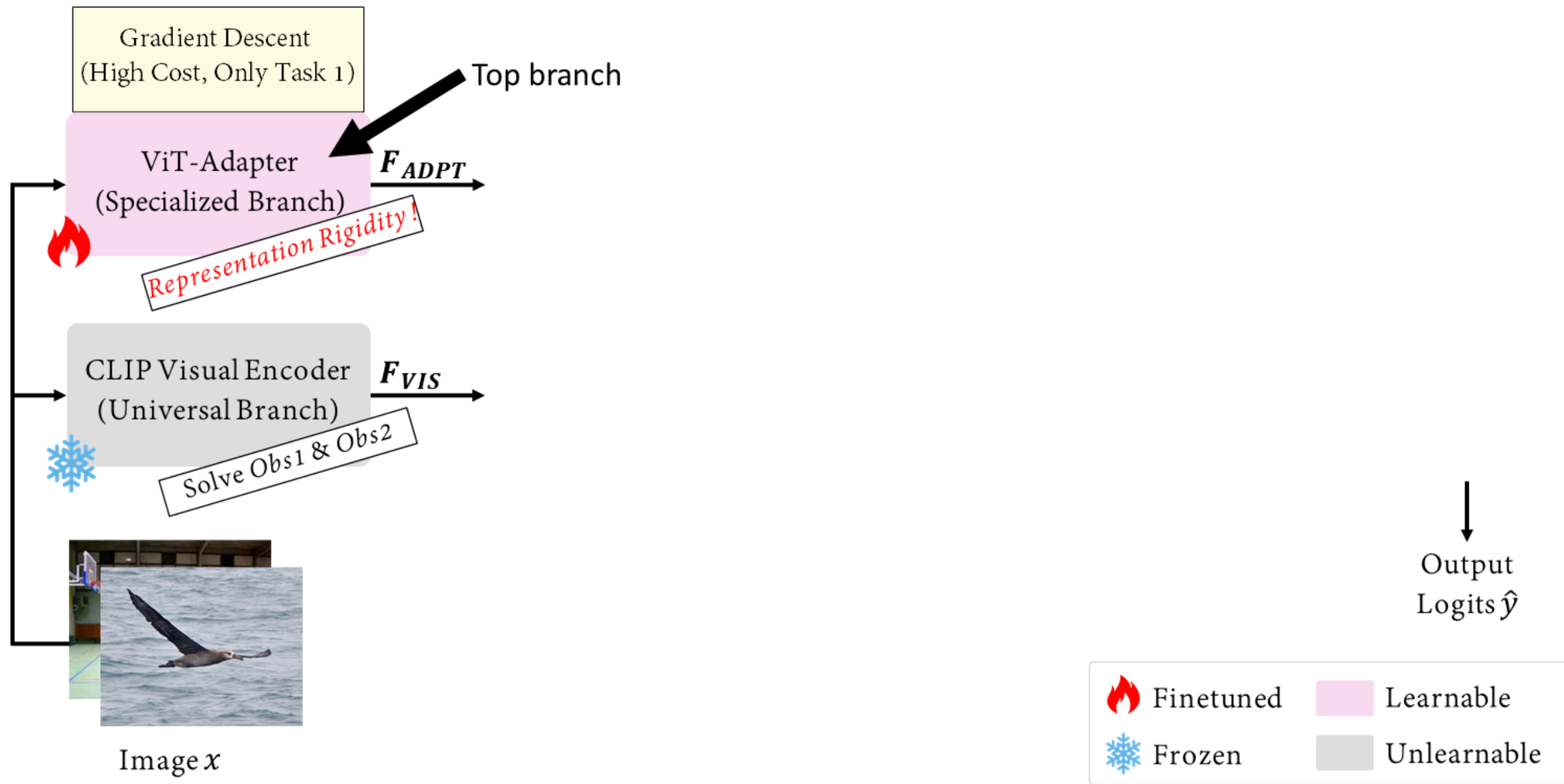
Complementary  
Features

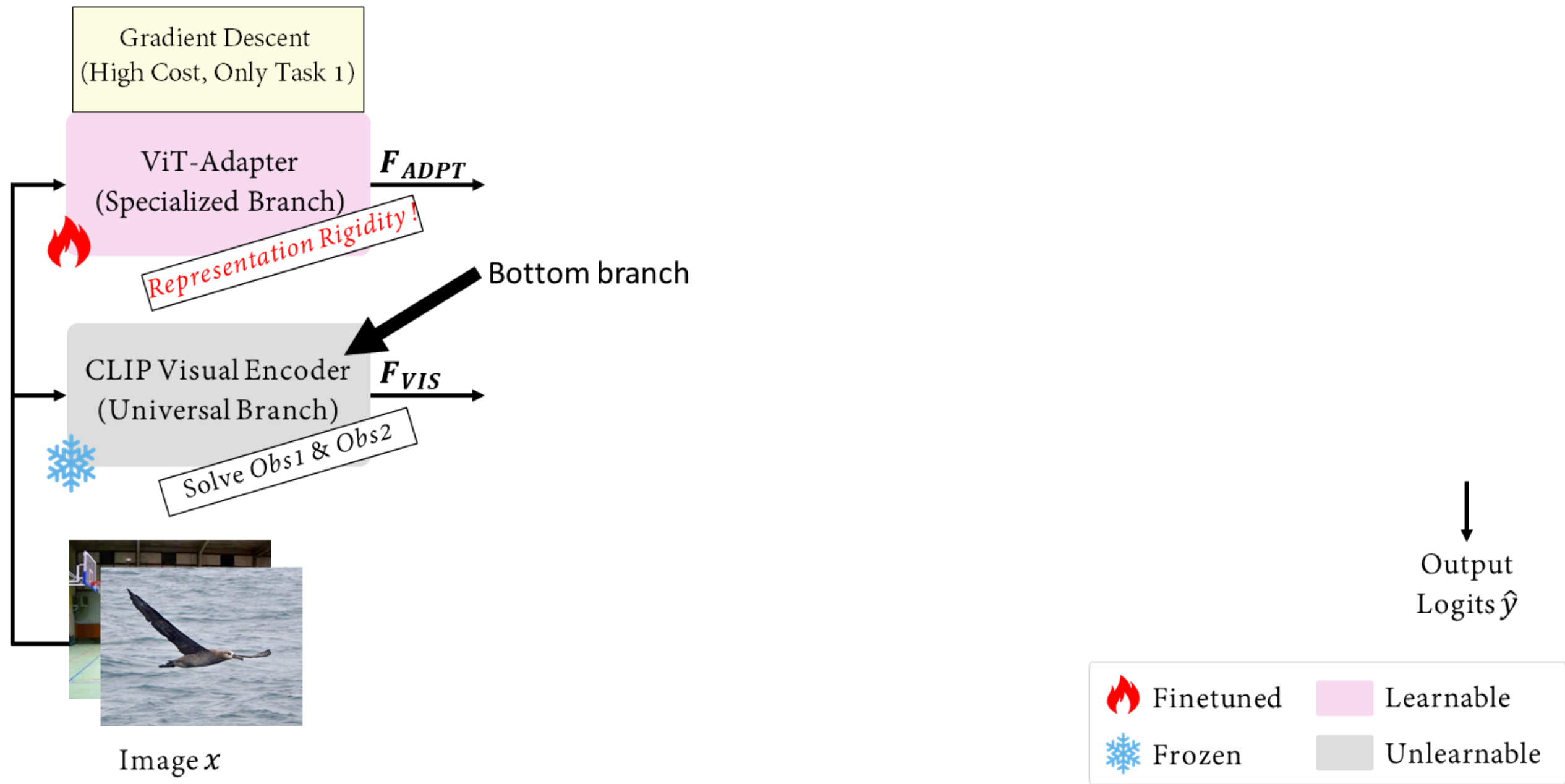
CLIP Subspace

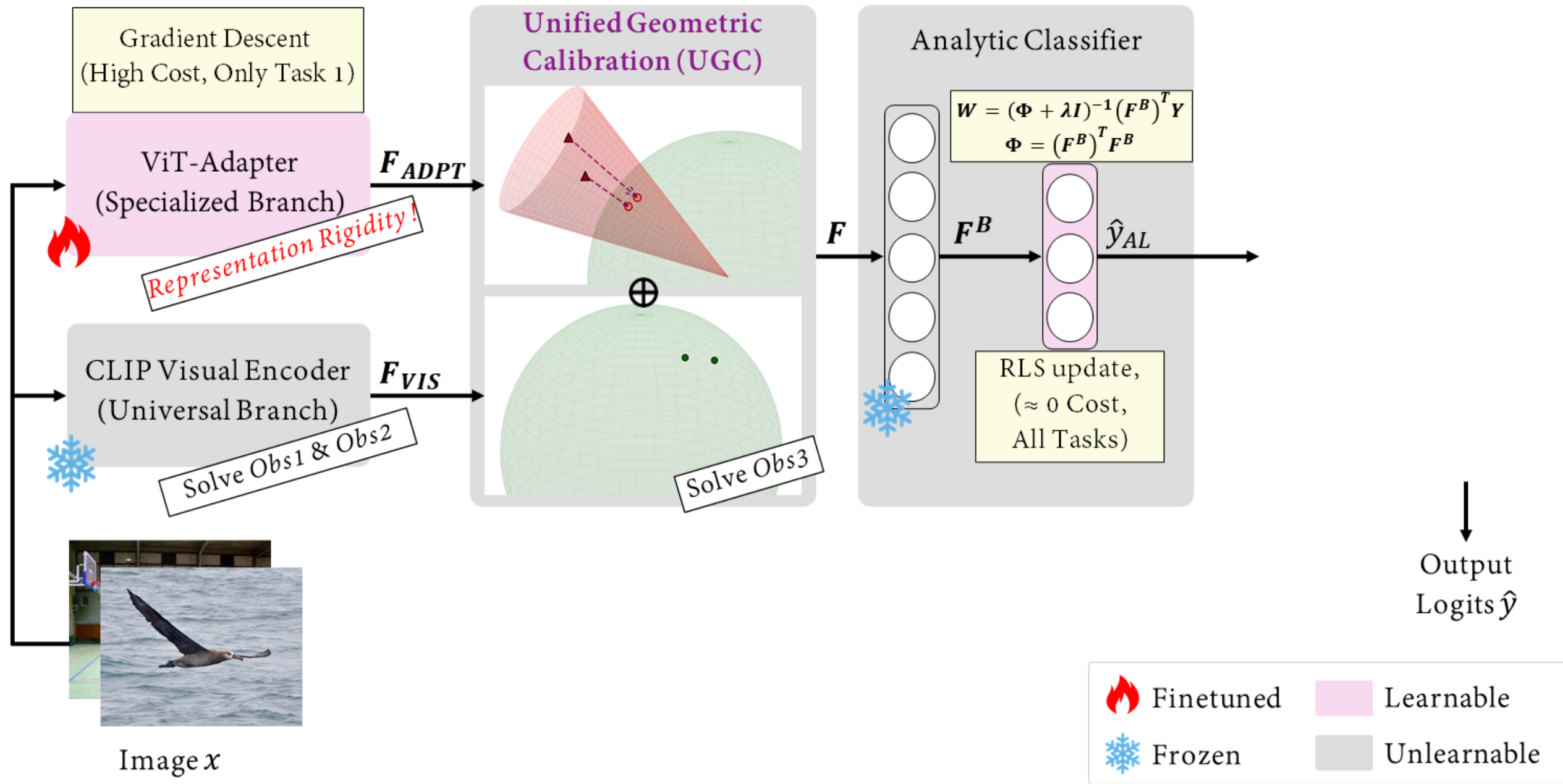
- ▲ Task 1 ViT-Adapter Features
- Task 1 Projected ViT-Adapter Features
- Task 1 CLIP Features
- Task 2 CLIP Features

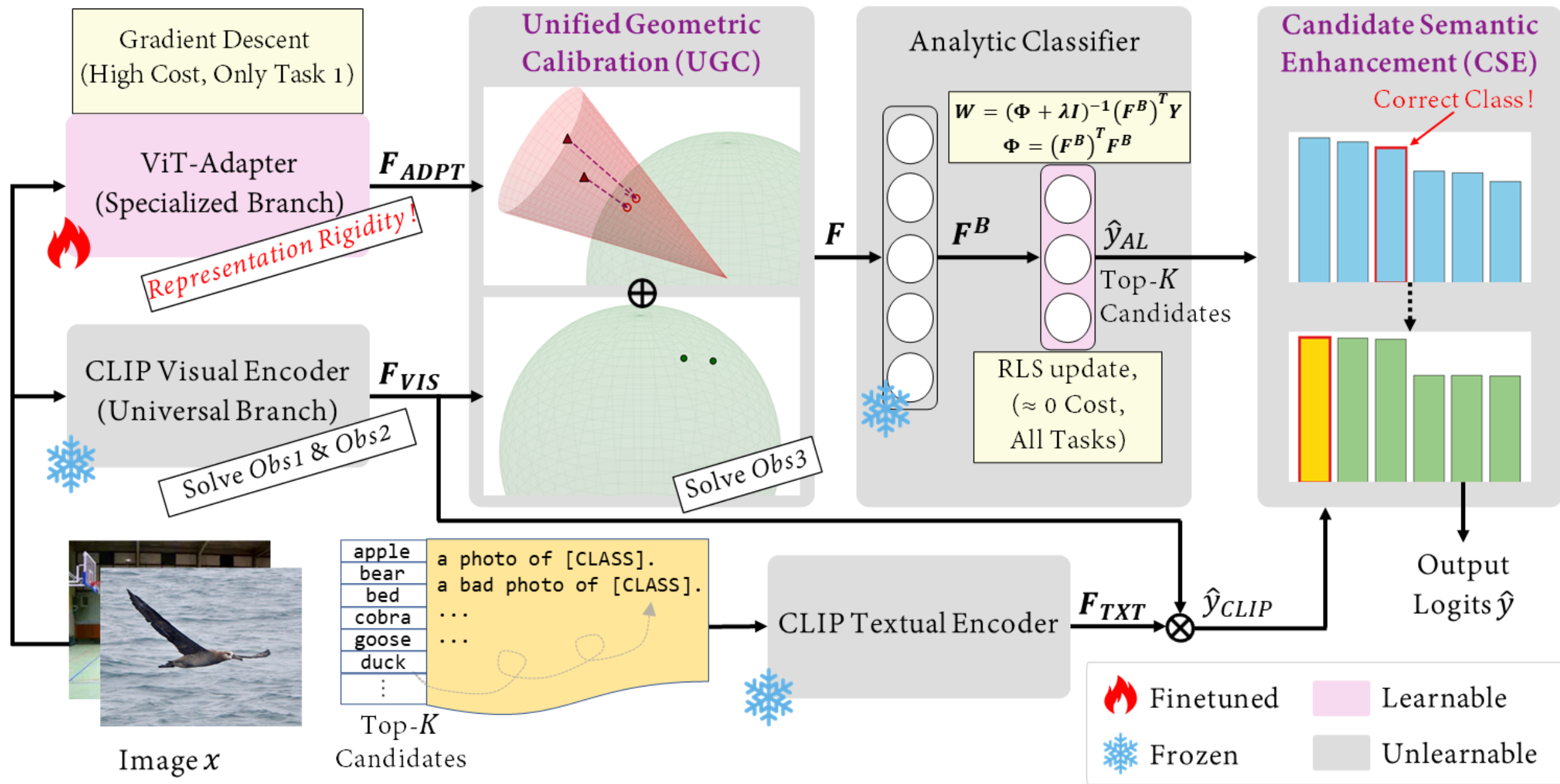
Obs1 (Collapse in long sequences):  
The ViT-Adapter subspace is narrow  
Obs2 (Single branch conflict):  
Shape of ViT-Adapter/CLIP subspace  
Obs3 (CLIP geometric misalignment):  
Subspace gap & topological difference

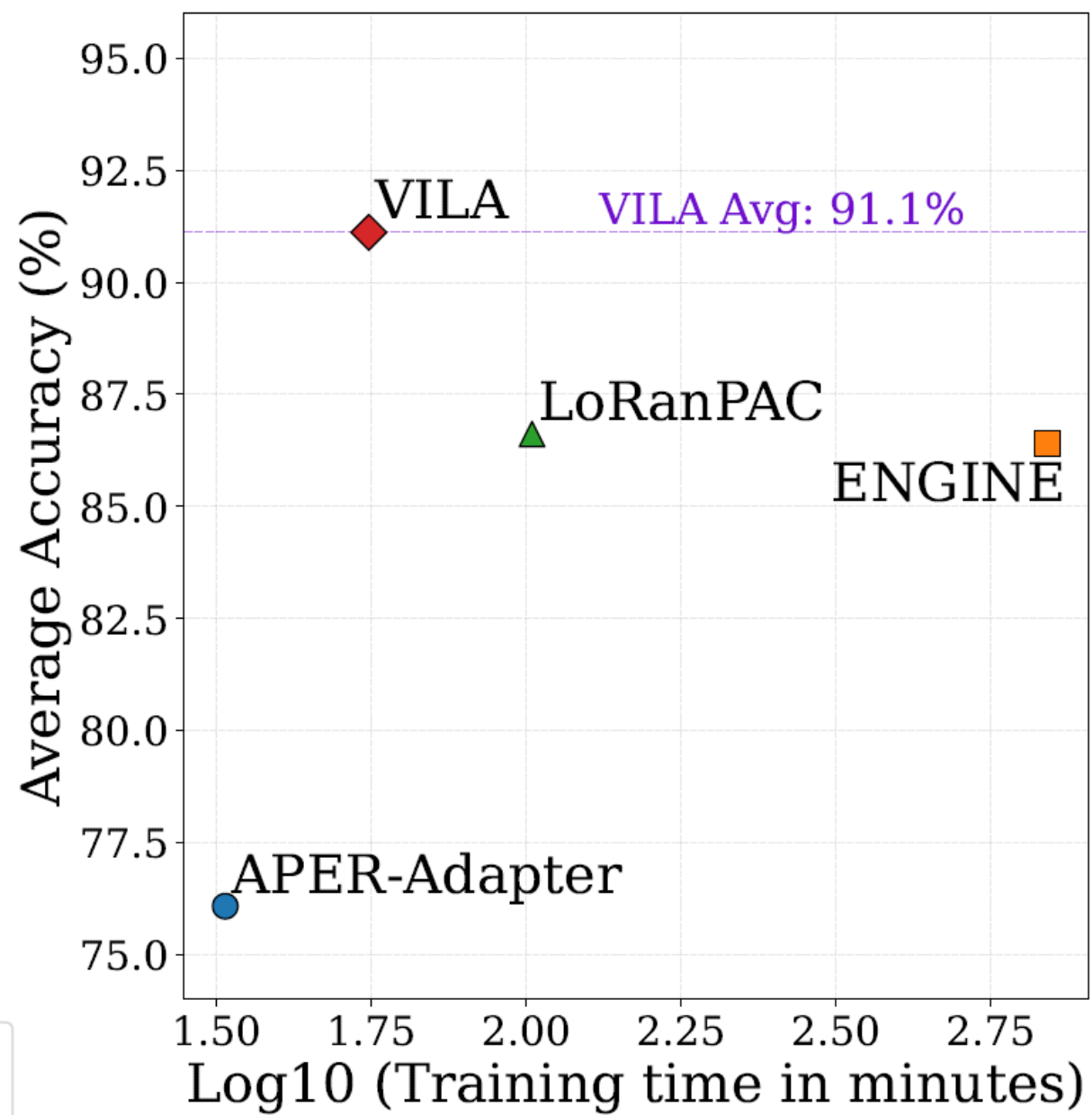
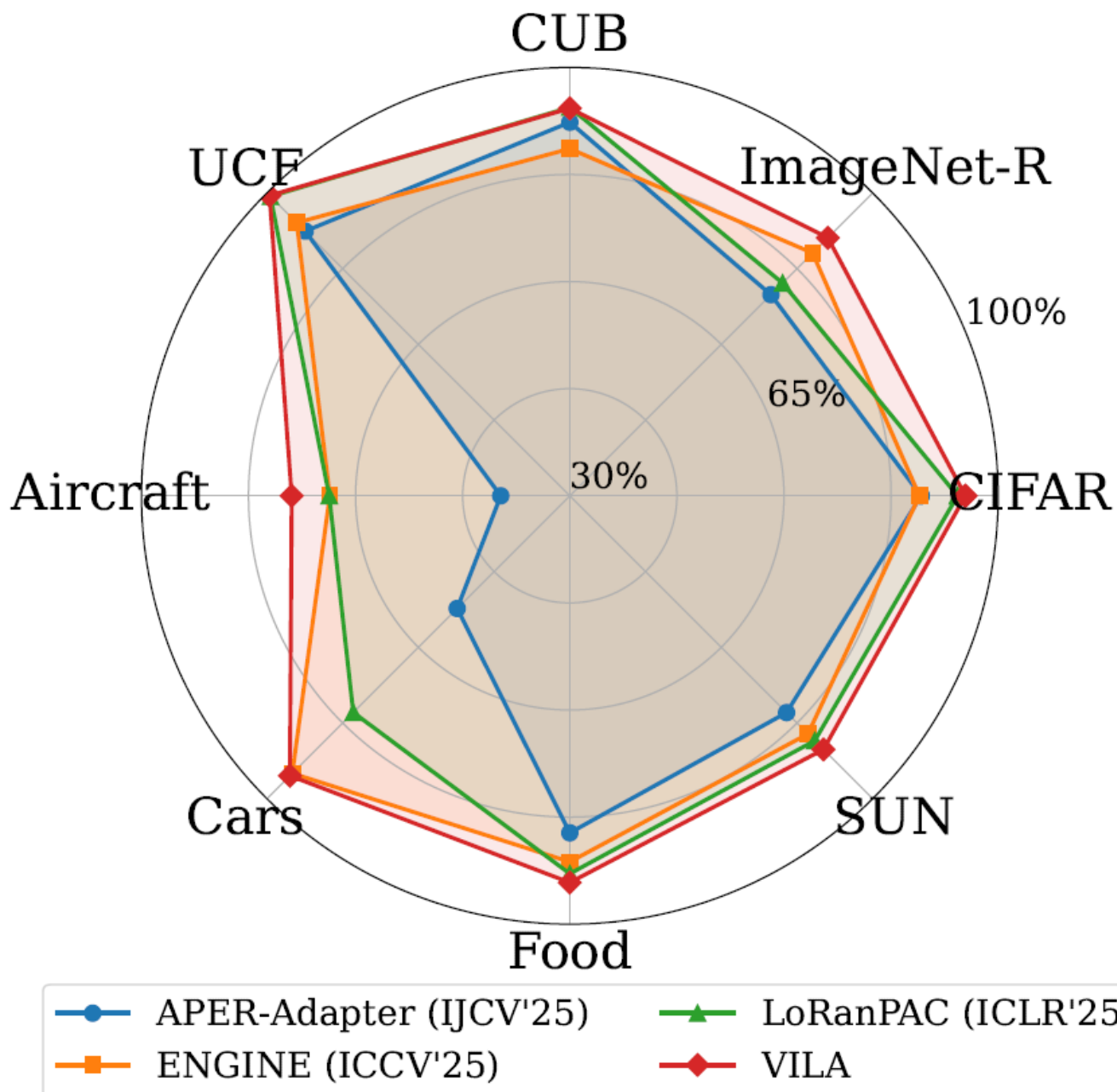












VILA breaks representation rigidity in analytic class-incremental learning via dual calibration: aligning task-adapted features with a universal anchor and rectifying decisions via semantic priors, achieving leading accuracy and extreme efficiency.



<https://byzhaoai.github.io/>

Binyu-Zhao@outlook.com