Plug-and-Play Compositionality for Boosting Continual Learning with Foundation Models

Weiduo Liao[1,3,4], Fei Han[2], *Hisao Ishibuchi*[3], *Qingfu Zhang*[4], *Ying Wei*[1]

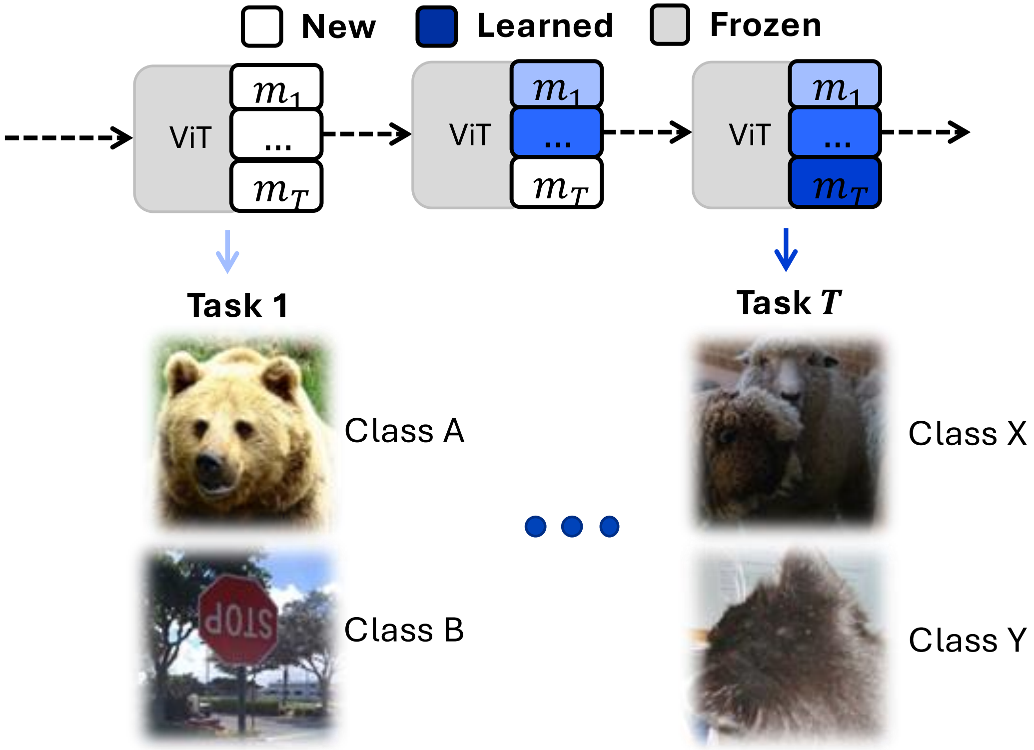
Date:

4/21/2026

1 2 香港城市大學 City University of Hong Kong 3 4

Continual Learning with Foundation Models (CL with FMs)

Sequential Class-Incremental Learning (CIL)

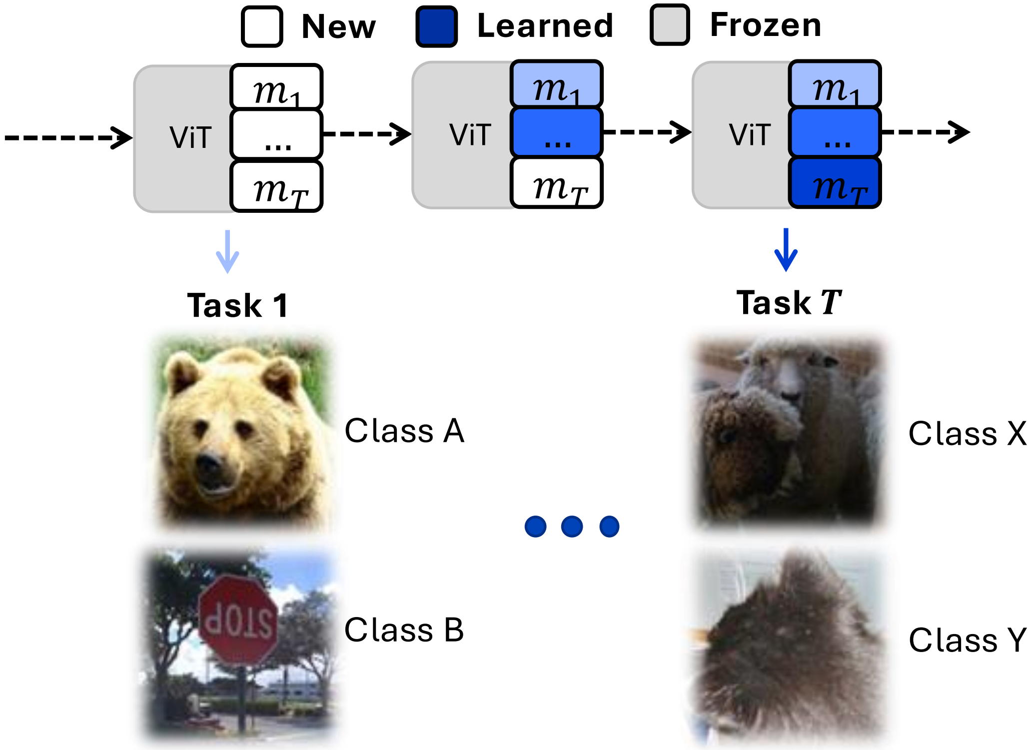


Continual Learning with Foundation Models (CL with FMs)

Sequential Class-Incremental Learning (CIL)

Why use FMs for CL?

- Powerful pre-trained backbones

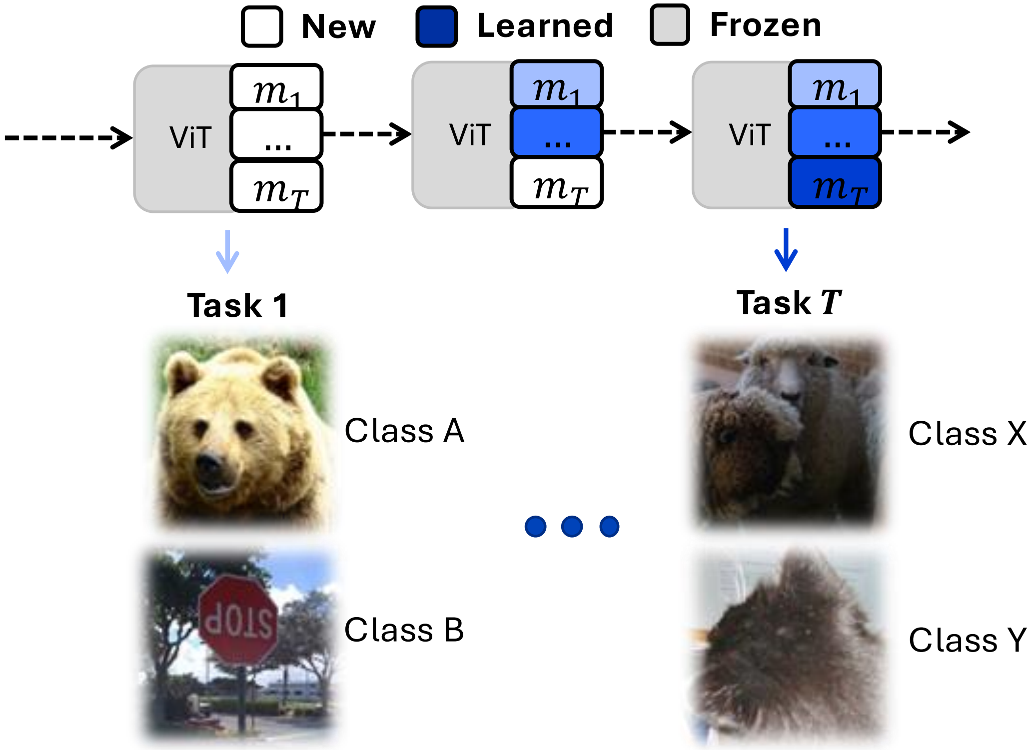


Continual Learning with Foundation Models (CL with FMs)

Sequential Class-Incremental Learning (CIL)

Why use FMs for CL?

- Powerful pre-trained backbones



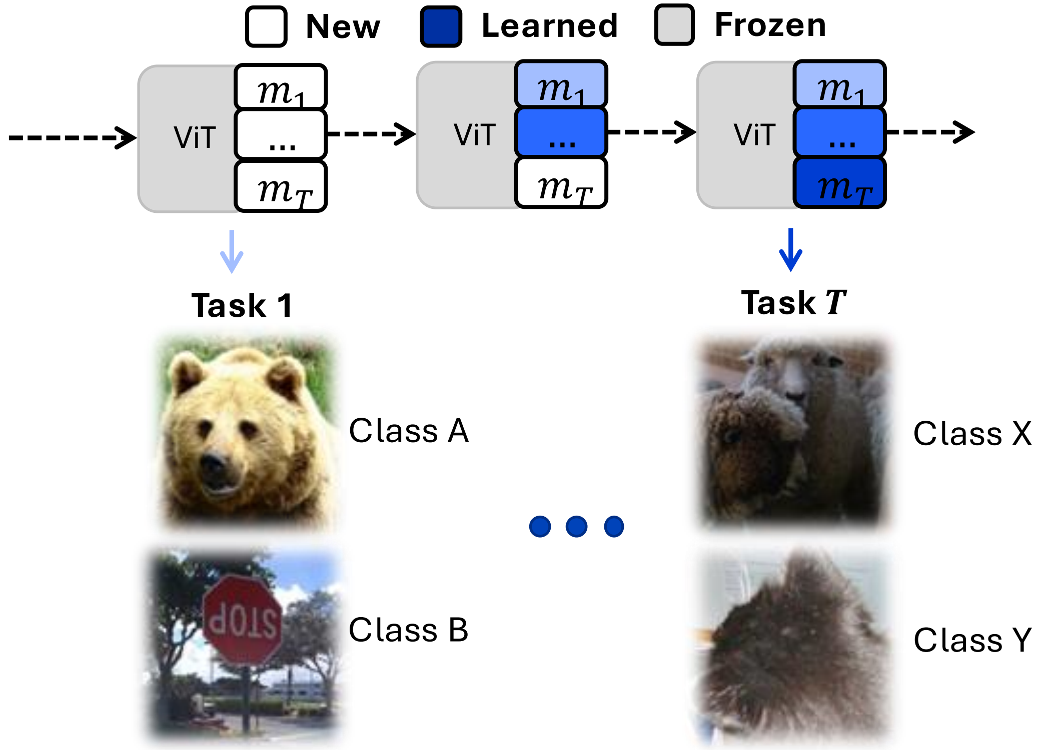
Why is forgetting still a critical issue?

Continual Learning with Foundation Models (CL with FMs)

Sequential Class-Incremental Learning (CIL)

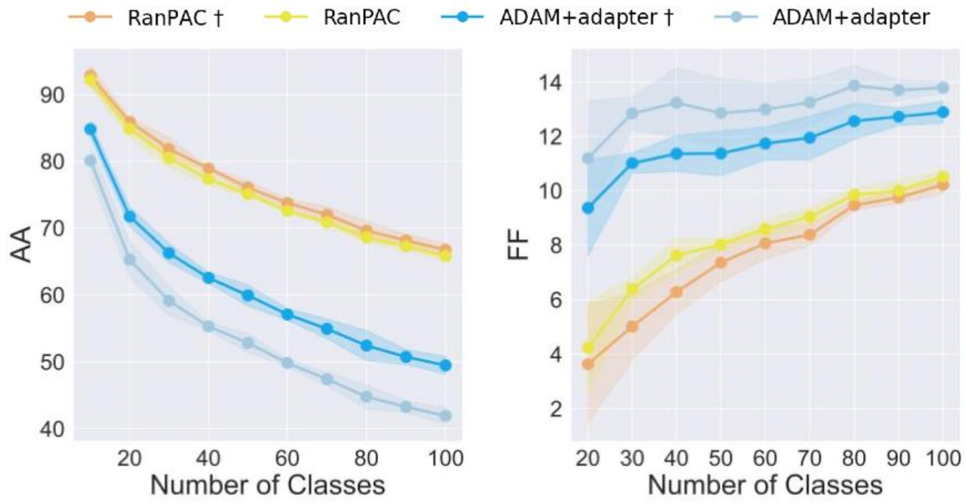
Why use FMs for CL?

- Powerful pre-trained backbones



Why is forgetting still a critical issue?

- Performance drop as classes increase
- Catastrophic forgetting of old tasks



How Do Current Methods Fight Forgetting?

Methods	Strategy	Representative Works
Prompt-based	Tuning task-specific prompt tokens without fine-tuning the frozen backbone	CPrompt[1], CODA-Prompt[2]
Representation-based	Leveraging backbone features with class prototype-based classifiers	ADAM+adapter[3], RanPAC[4]
Model-mixture	Using hybrid techniques like model fusion or ensembling for robust prediction	CoFiMA[5]
Rehearsal-based	Reviewing past samples or auxiliary supervision to address class imbalance	FOSTER[6], DER[7], MEMO[8]

[1] Z. Gao et al. Consistent prompting for rehearsal-free continual learning. CVPR 2024
[2] J.S. Smith et al. Coda-prompt: Continual decomposed attention-based prompting for rehearsal-free continual learning. CVPR 2023
[3] D.W. Zhou et al. Revisiting class-incremental learning with pre-trained models: Generalizability and adaptivity are all you need. IJCV 2025.
[4] M.D. McDonnell et al. Ranpac: Random projections and pre-trained models for continual learning. NeurIPS 2023
[5] I.E. Marouf et al. Weighted ensemble models are strong continual learners. ECCV 2024.
[6] F.Y. Wang et al. Foster: Feature boosting and compression for class-incremental learning. ECCV 2022.
[7] S. Yan et al. Der: Dynamically expandable representation for class incremental learning. CVPR 2021.
[8] D.W. Zhou et al. A model or 603 exemplars: Towards memory-efficient class-incremental learning. ICLR 2023.

What is missing in current continual learning methods?

Task 1

Class A

Class B

* Images generated by Gemini Nano Banana.

What is missing in current continual learning methods?

Task 1

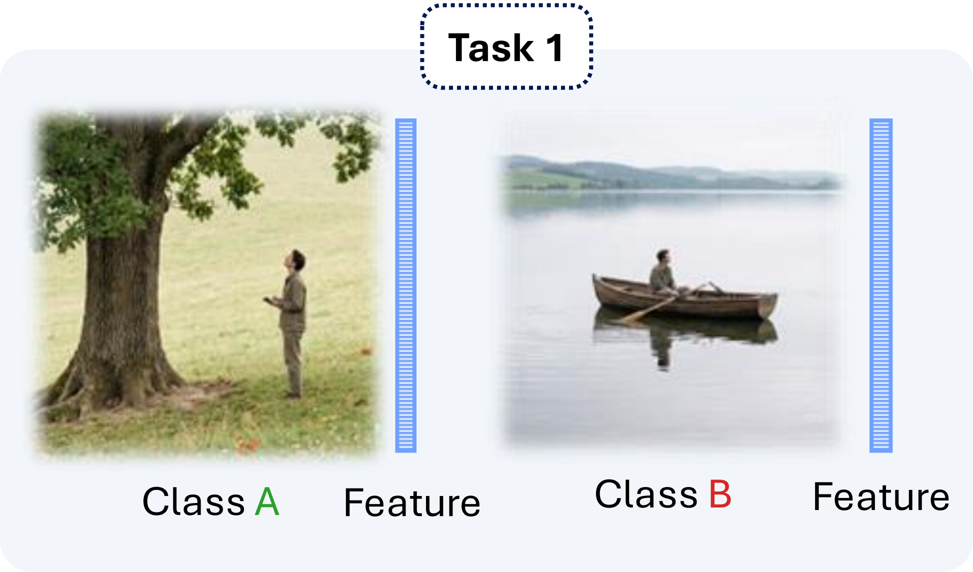
Class A

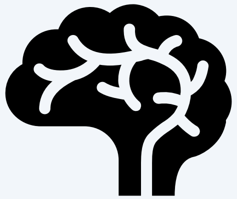
Class B

Rich Knowledge Base

* Images generated by Gemini Nano Banana.

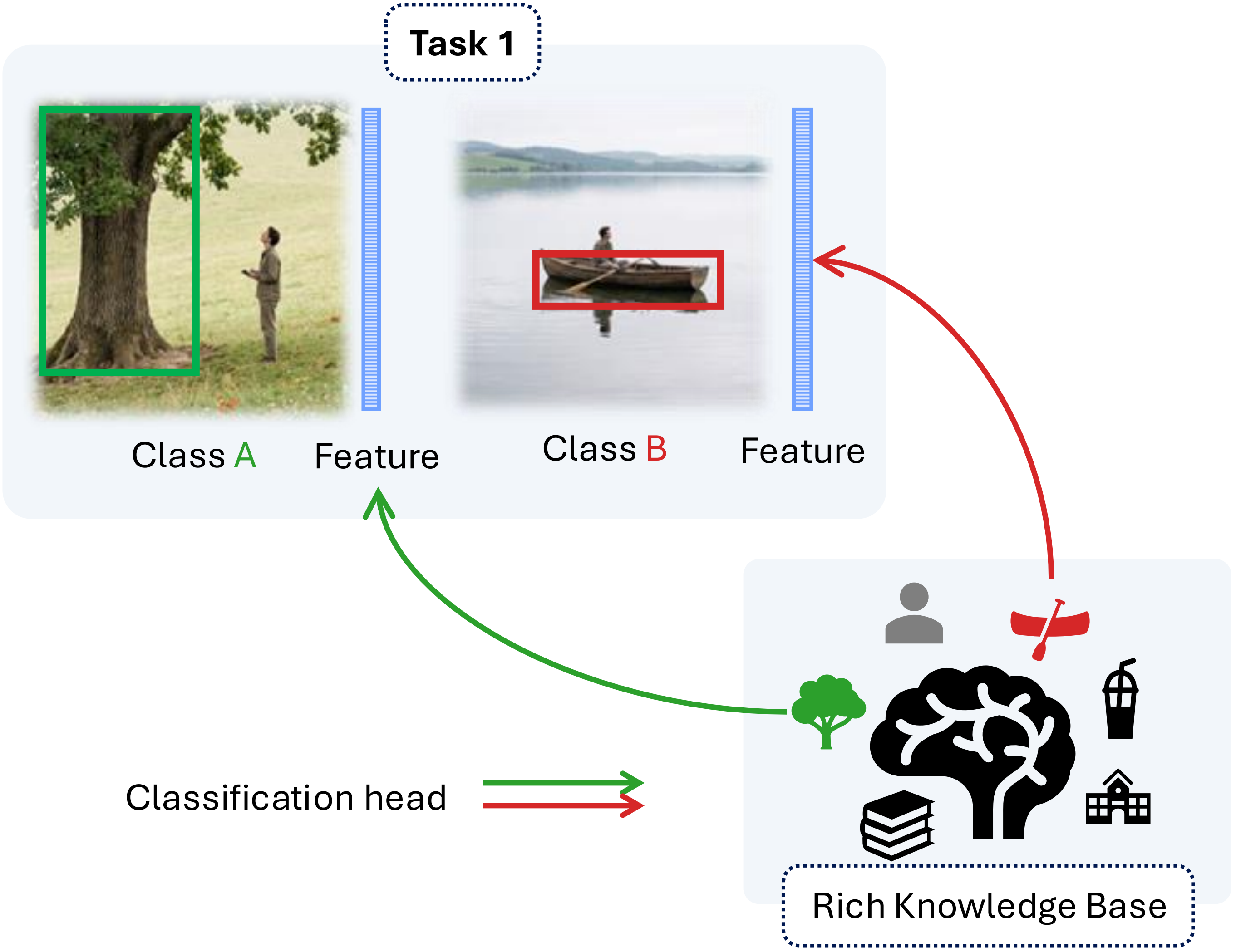
What is missing in current continual learning methods?





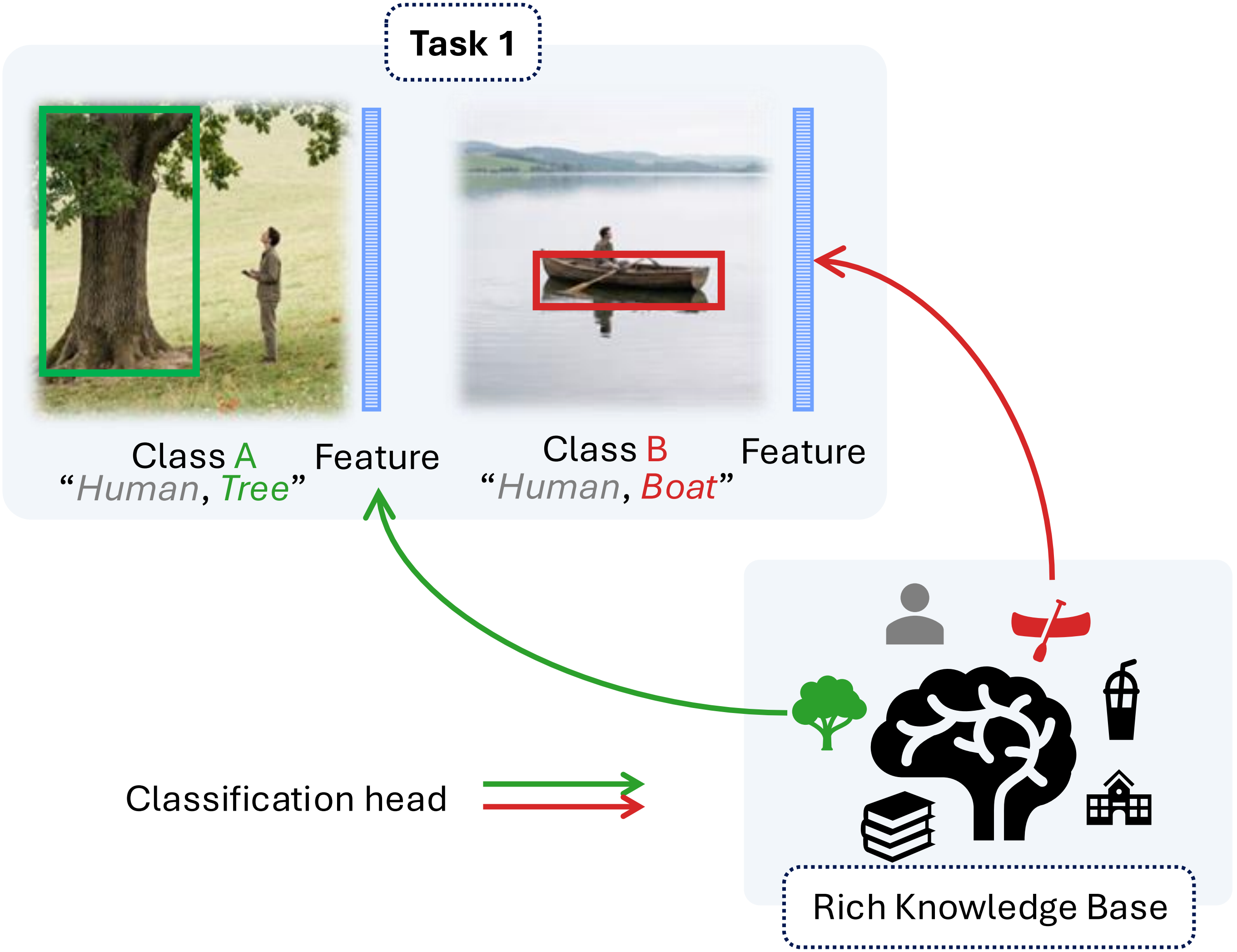
* Images generated by Gemini Nano Banana.

What is missing in current continual learning methods?



* Images generated by Gemini Nano Banana.

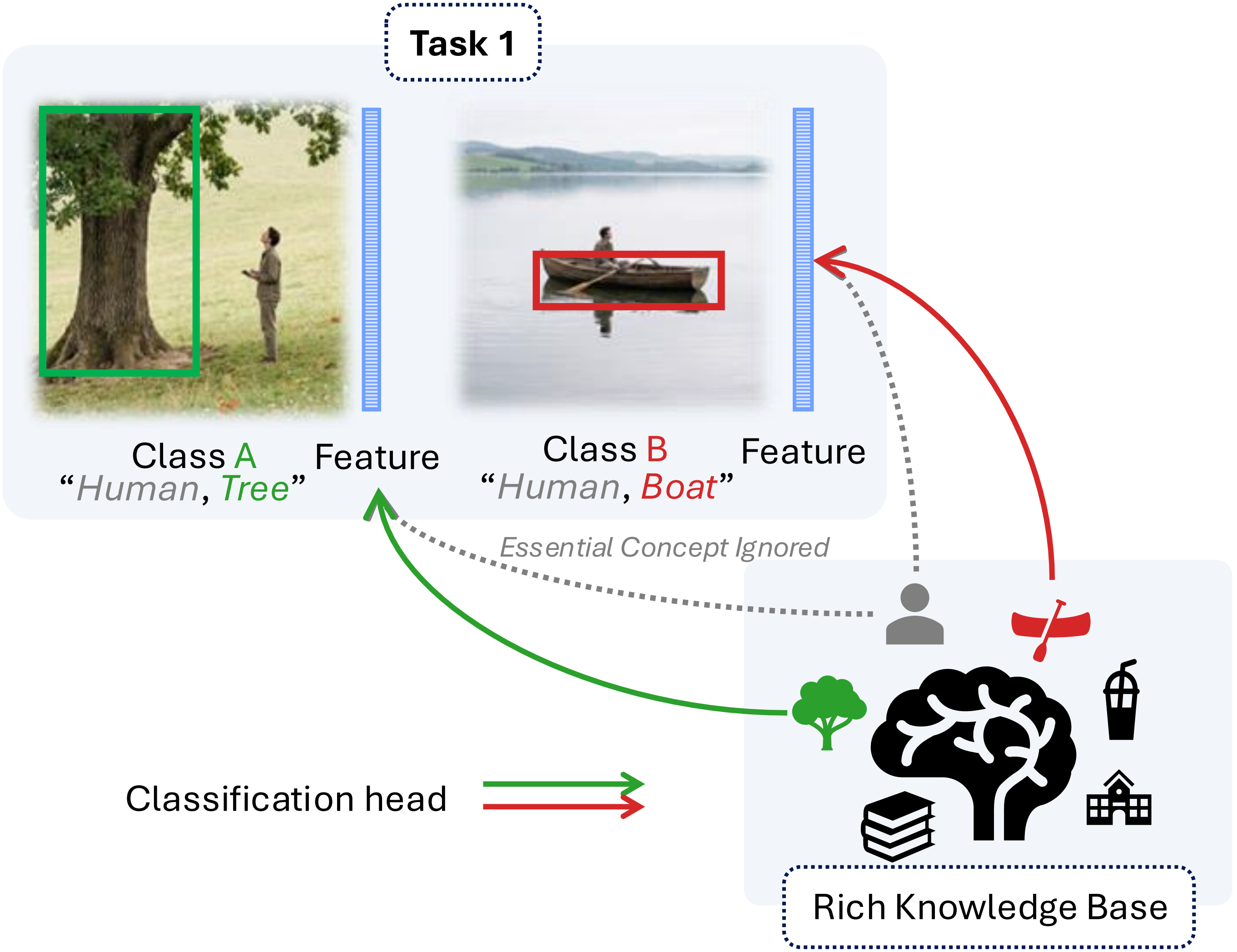
What is missing in current continual learning methods?



* Images generated by Gemini Nano Banana.

What is missing in current continual learning methods?

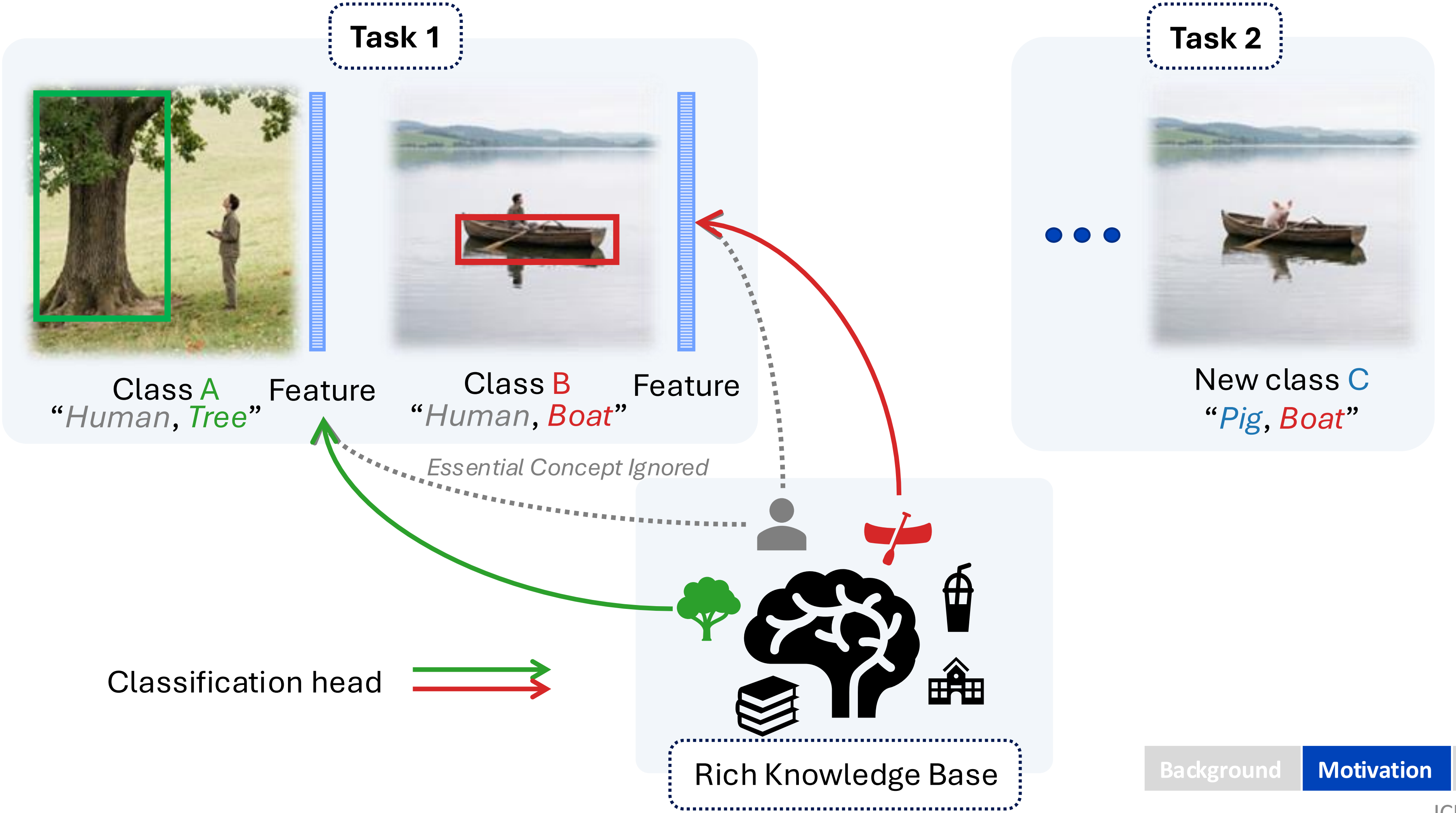
High-dimensional un-structural feature space causes neglect of essential concepts (e.g., *Human*) for easy comparison paths (e.g., *Tree* *v.s.* *Boat*).



* Images generated by Gemini Nano Banana.

What is missing in current continual learning methods?

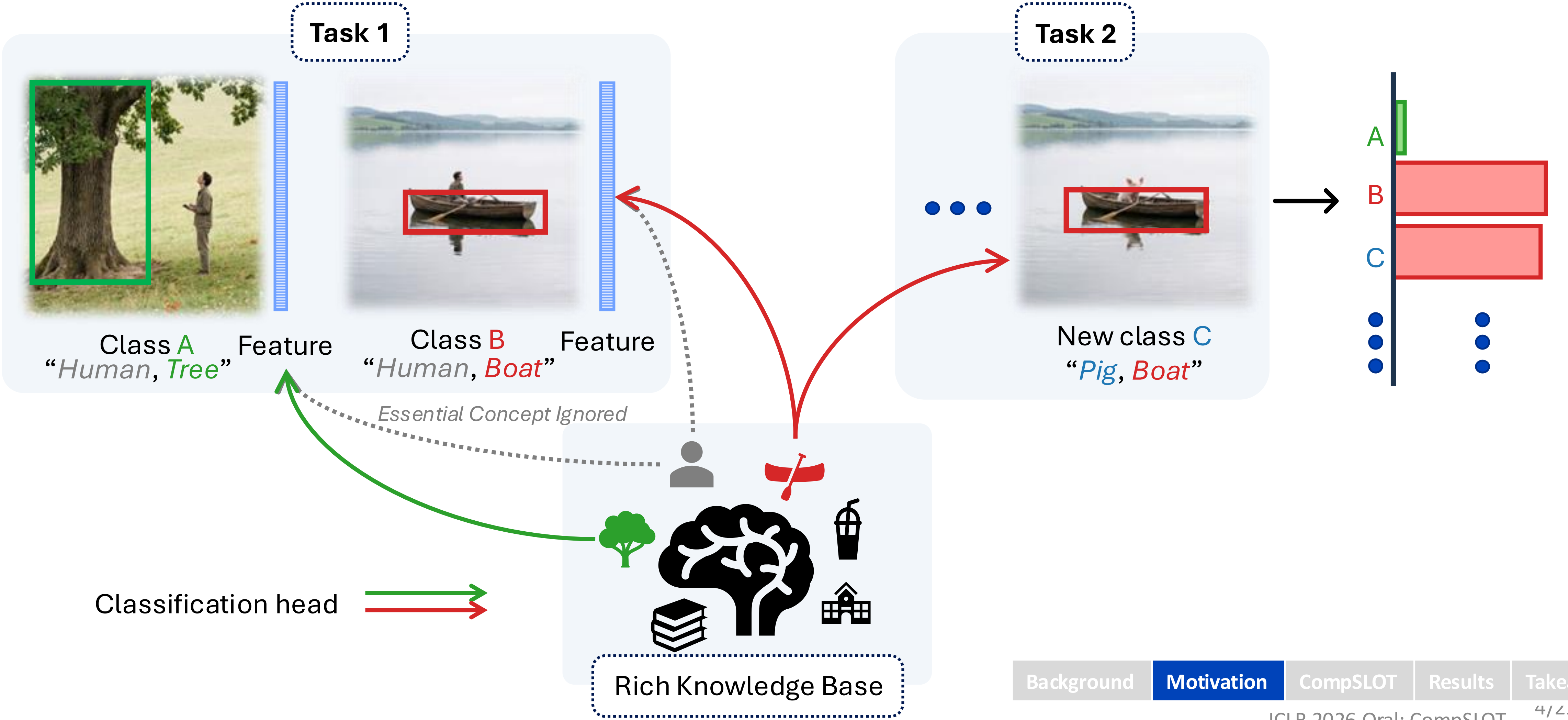
High-dimensional un-structural feature space causes neglect of essential concepts (e.g., *Human*) for easy comparison paths (e.g., *Tree* *v.s.* *Boat*).



* Images generated by Gemini Nano Banana.

What is missing in current continual learning methods?

High-dimensional un-structural feature space causes neglect of essential concepts (e.g., *Human*) for easy comparison paths (e.g., *Tree* *v.s.* *Boat*).



* Images generated by Gemini Nano Banana.

From High-Dim Un-Structural Features to Structural Low-Dim Concepts

Pre-defined structural concept space
{*Human*, *Tree*, *Boat*, *Pig*}

Task 1

Class A

Class B

Task 2

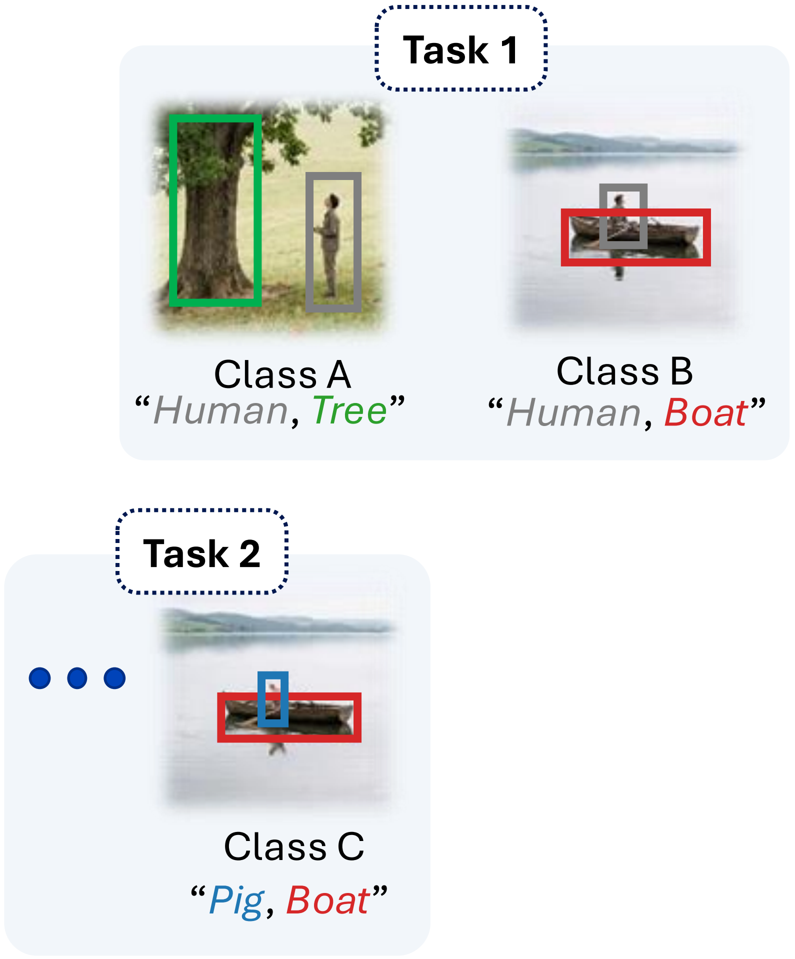
• • •

Class C

From High-Dim Un-Structural Features to Structural Low-Dim Concepts

Pre-defined structural concept space
{*Human*, *Tree*, *Boat*, *Pig*}

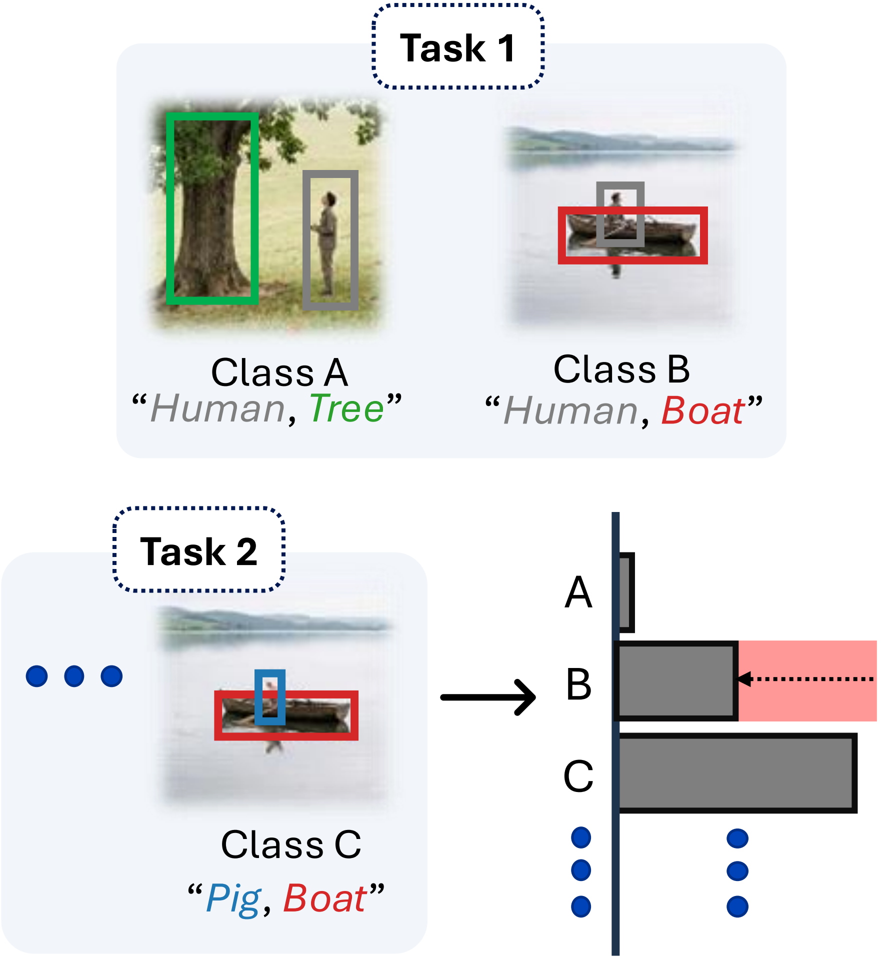
- Essential ***Human*** becomes prioritized and unavoidable.



From High-Dim Un-Structural Features to Structural Low-Dim Concepts

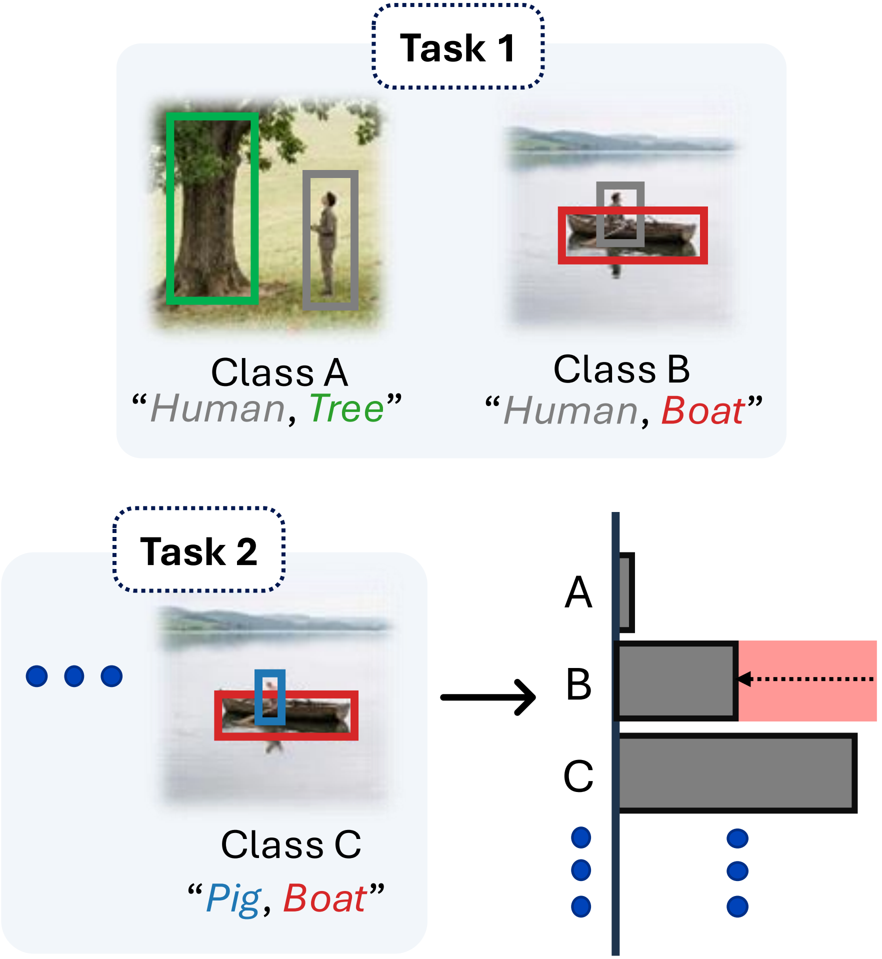
Pre-defined structural concept space
{*Human*, *Tree*, *Boat*, *Pig*}

- Essential ***Human*** becomes prioritized and unavoidable.



Pre-defined structural concept space
{*Human*, *Tree*, *Boat*, *Pig*}

- Essential ***Human*** becomes prioritized and unavoidable.

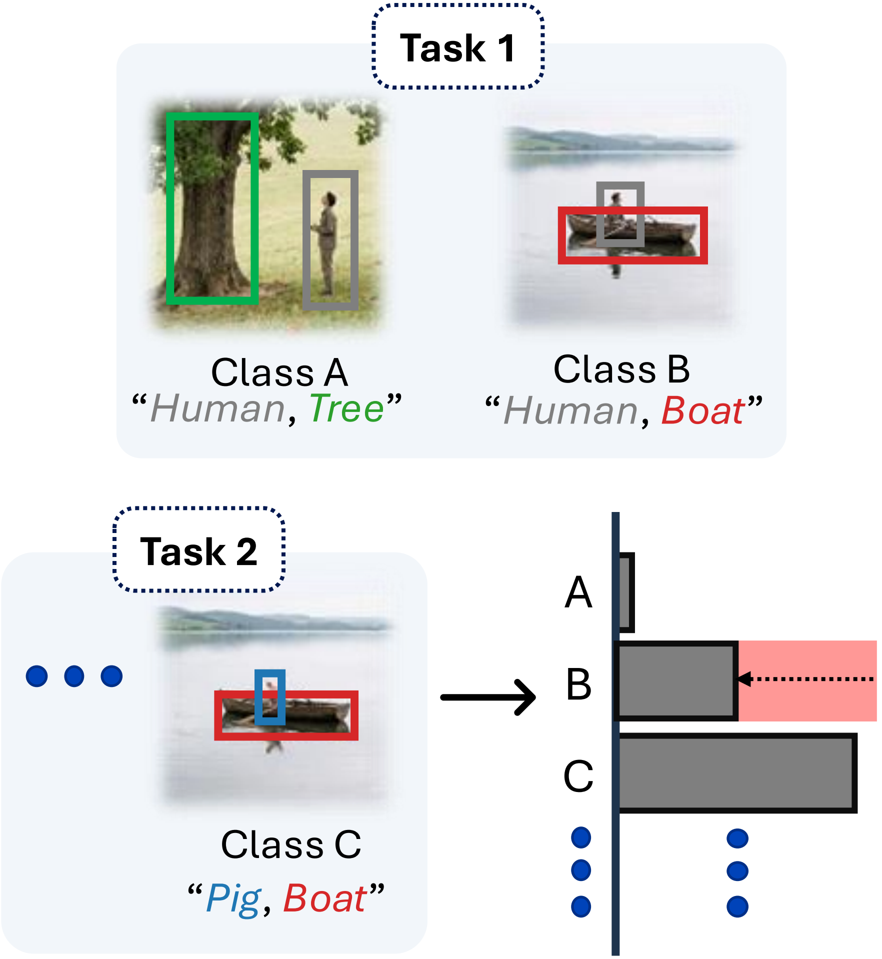


Challenge 1: Is there inherent concept awareness in Foundation Models?

From High-Dim Un-Structural Features to Structural Low-Dim Concepts

Pre-defined structural concept space
{*Human*, *Tree*, *Boat*, *Pig*}

- Essential ***Human*** becomes prioritized and unavoidable.



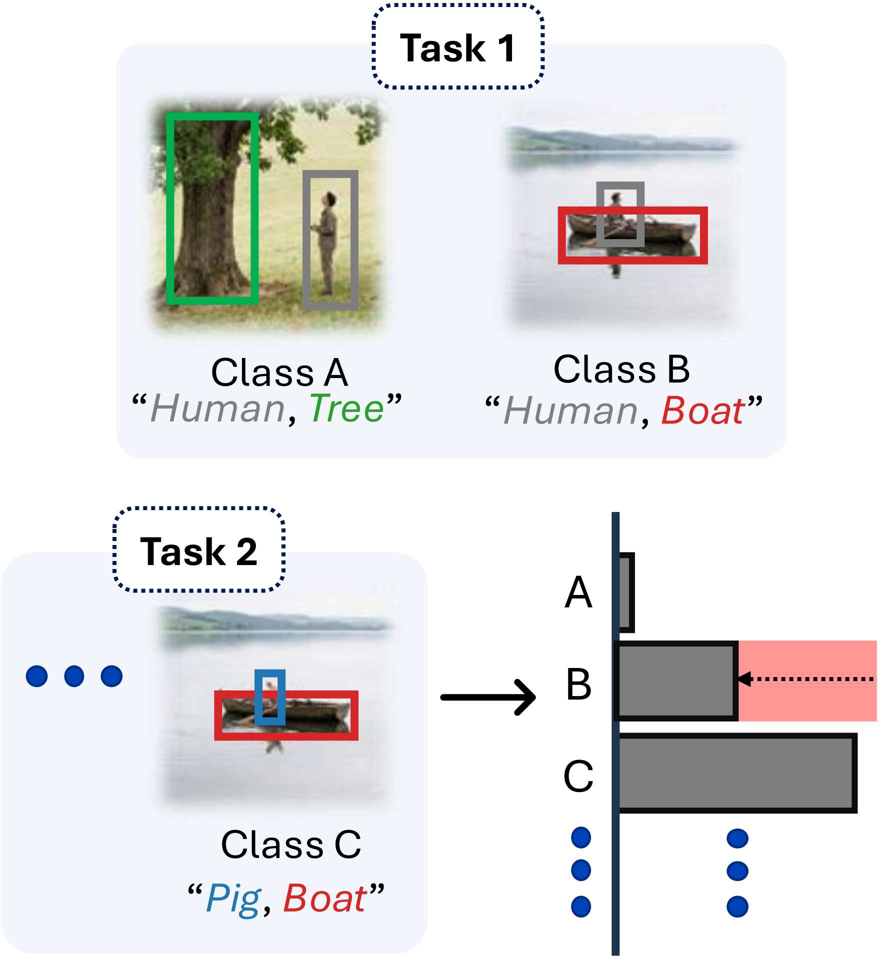
Challenge 1: Is there inherent concept awareness in Foundation Models?

- Relying solely on the FM's internal semantic features to recognize important concepts
- For the sake of fairness

From High-Dim Un-Structural Features to Structural Low-Dim Concepts

Pre-defined structural concept space
{*Human*, *Tree*, *Boat*, *Pig*}

- Essential ***Human*** becomes prioritized and unavoidable.



Challenge 1: Is there inherent concept awareness in Foundation Models?

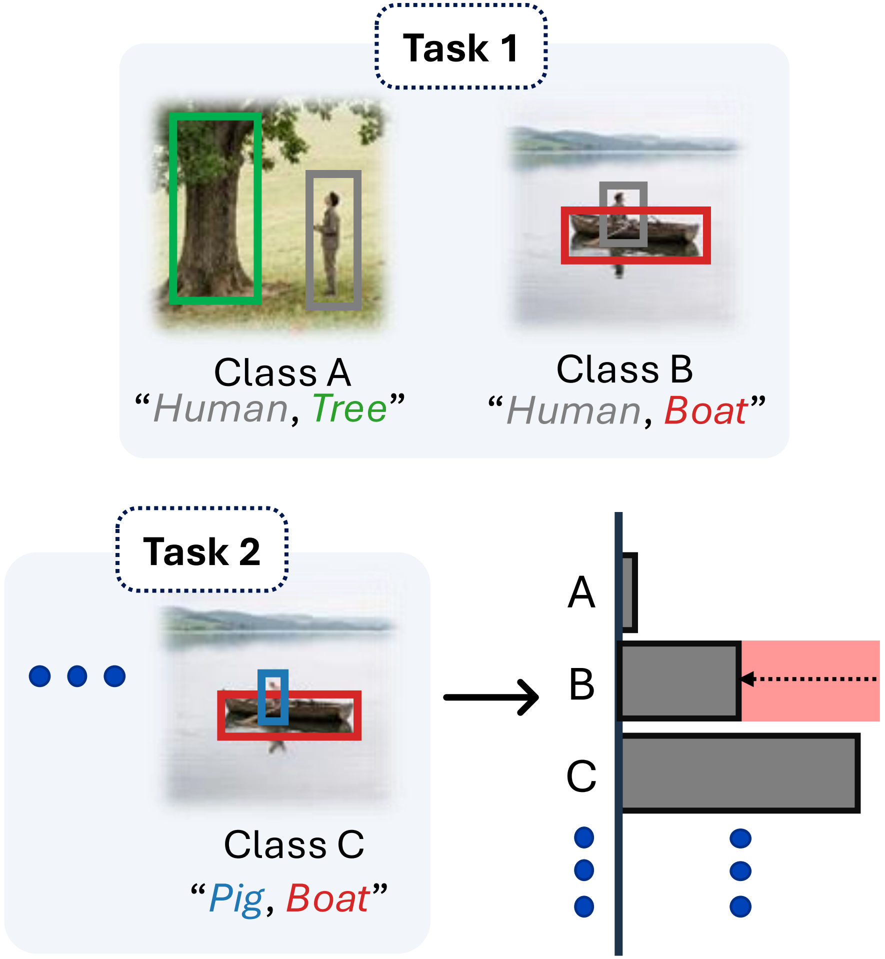
- Relying solely on the FM's internal semantic features to recognize important concepts
- For the sake of fairness

Challenge 2: How to establish cross-class concept grounding for a wide range of CL algorithms?

From High-Dim Un-Structural Features to Structural Low-Dim Concepts

Pre-defined structural concept space
{*Human*, *Tree*, *Boat*, *Pig*}

- Essential ***Human*** becomes prioritized and unavoidable.



Challenge 1: Is there inherent concept awareness in Foundation Models?

- Relying solely on the FM's internal semantic features to recognize important concepts
- For the sake of fairness

Challenge 2: How to establish cross-class concept grounding for a wide range of CL algorithms?

- Building a lightweight, universal plug-in

Challenge 1: Is there inherent concept awareness in Foundation Model?

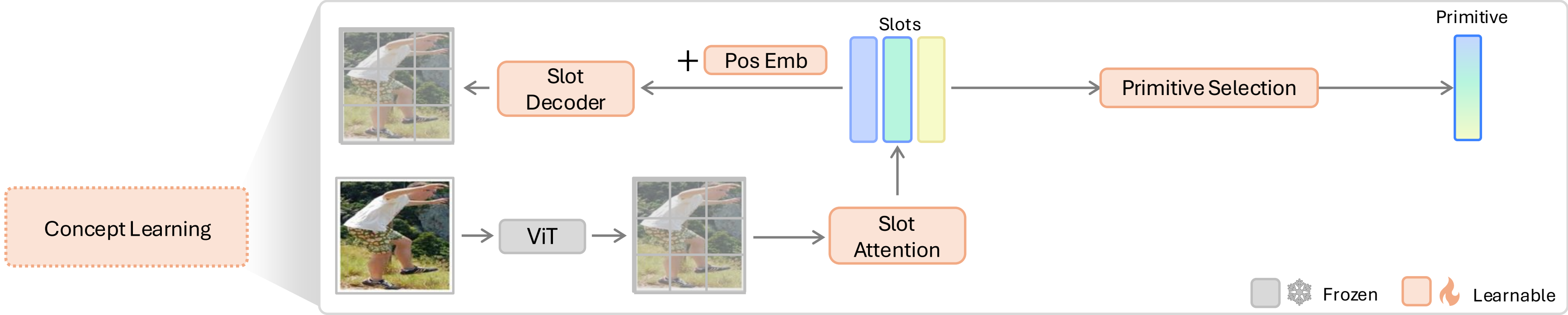
Challenge 2: How to establish cross-class concept grounding for a wide range of CL algorithms?

[1] F. Locatello, et al. Object-centric learning with slot attention. NeurIPS 2020.

Plug-and-Play CompSLOT

Challenge 1: Is there inherent concept awareness in Foundation Model?

- Stage 1: Concept learning via Slot Attention [1] and primitive selection



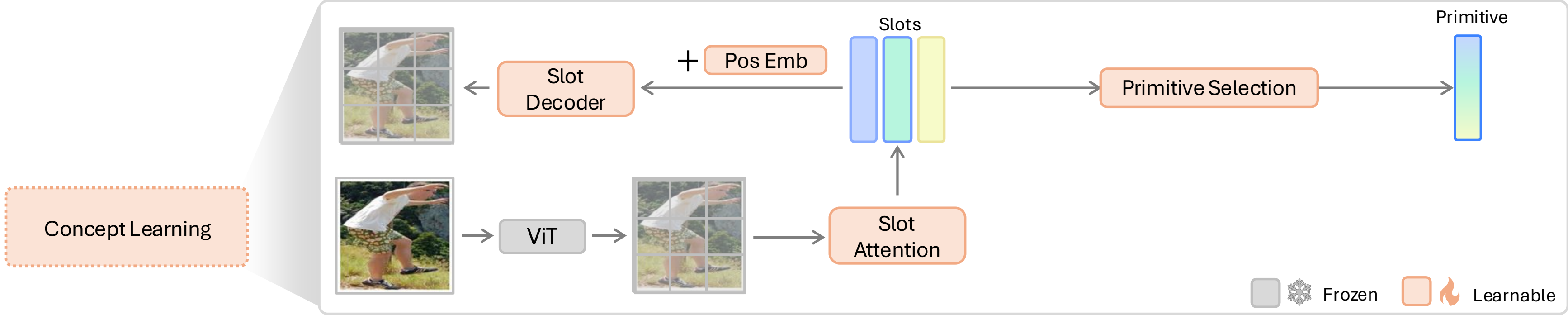
Challenge 2: How to establish cross-class concept grounding for a wide range of CL algorithms?

[1] F. Locatello, et al. Object-centric learning with slot attention. NeurIPS 2020.

Plug-and-Play CompSLOT

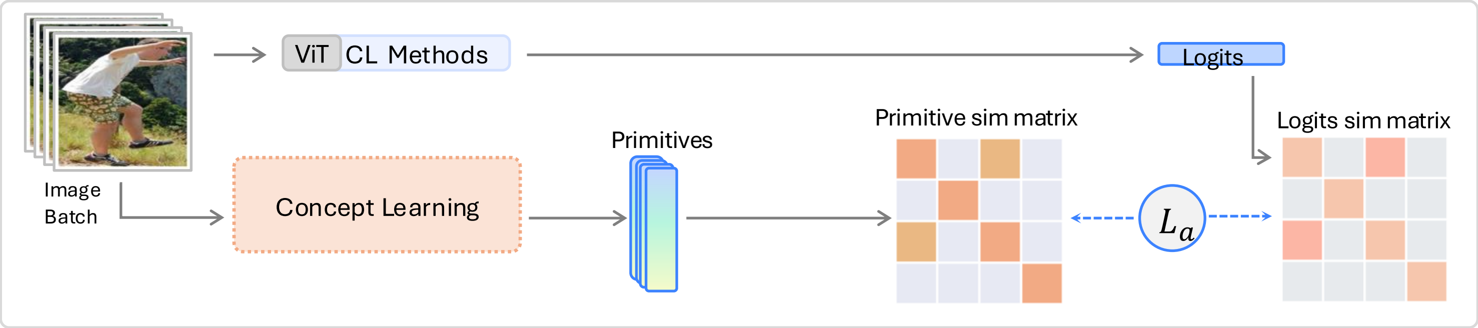
Challenge 1: Is there inherent concept awareness in Foundation Model?

- Stage 1: Concept learning via Slot Attention [1] and primitive selection



Challenge 2: How to establish cross-class concept grounding for a wide range of CL algorithms?

- Stage 2: Pair-wise distillation on logits using conceptual relationship



[1] F. Locatello, et al. Object-centric learning with slot attention. NeurIPS 2020.

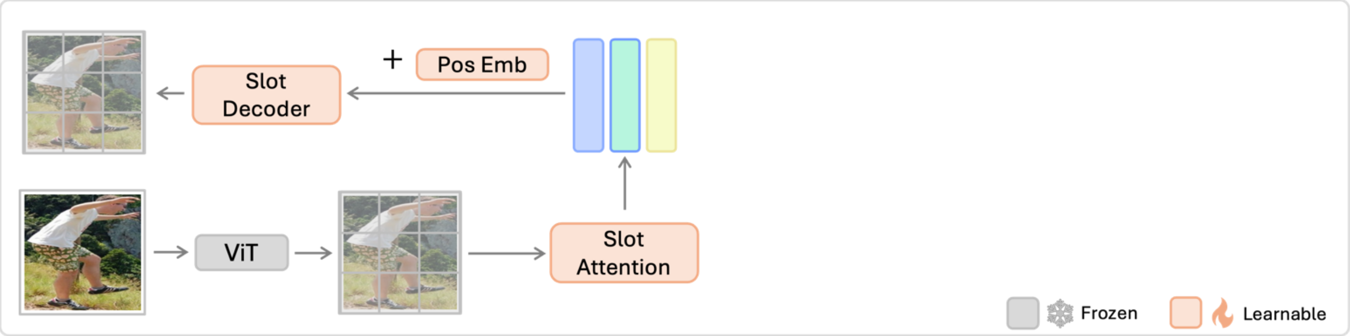
Stage 1: Concept Learning

Goal for challenge 1: Extracting conceptual image representations

Stage 1: Concept Learning

Goal for challenge 1: Extracting conceptual image representations

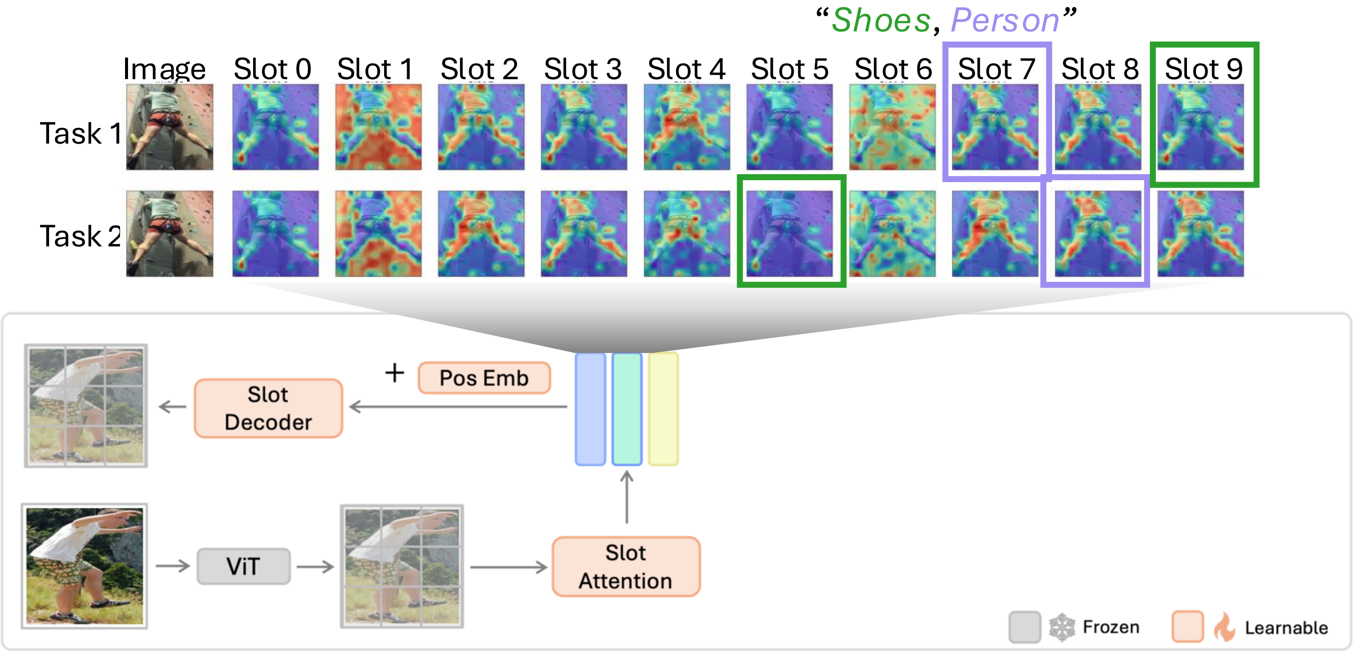
- Slot Attention: Decomposition into K independent slots



Stage 1: Concept Learning

Goal for challenge 1: Extracting conceptual image representations

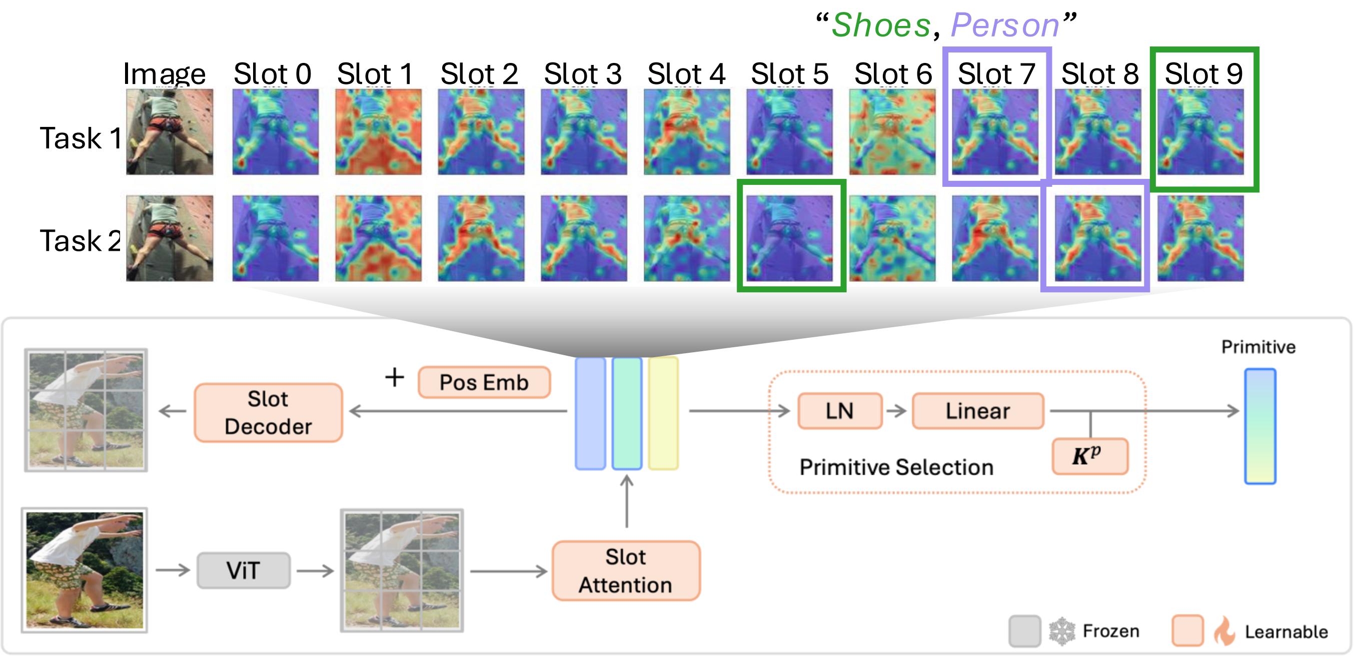
- Slot Attention: Decomposition into K independent slots



Stage 1: Concept Learning

Goal for challenge 1: Extracting conceptual image representations

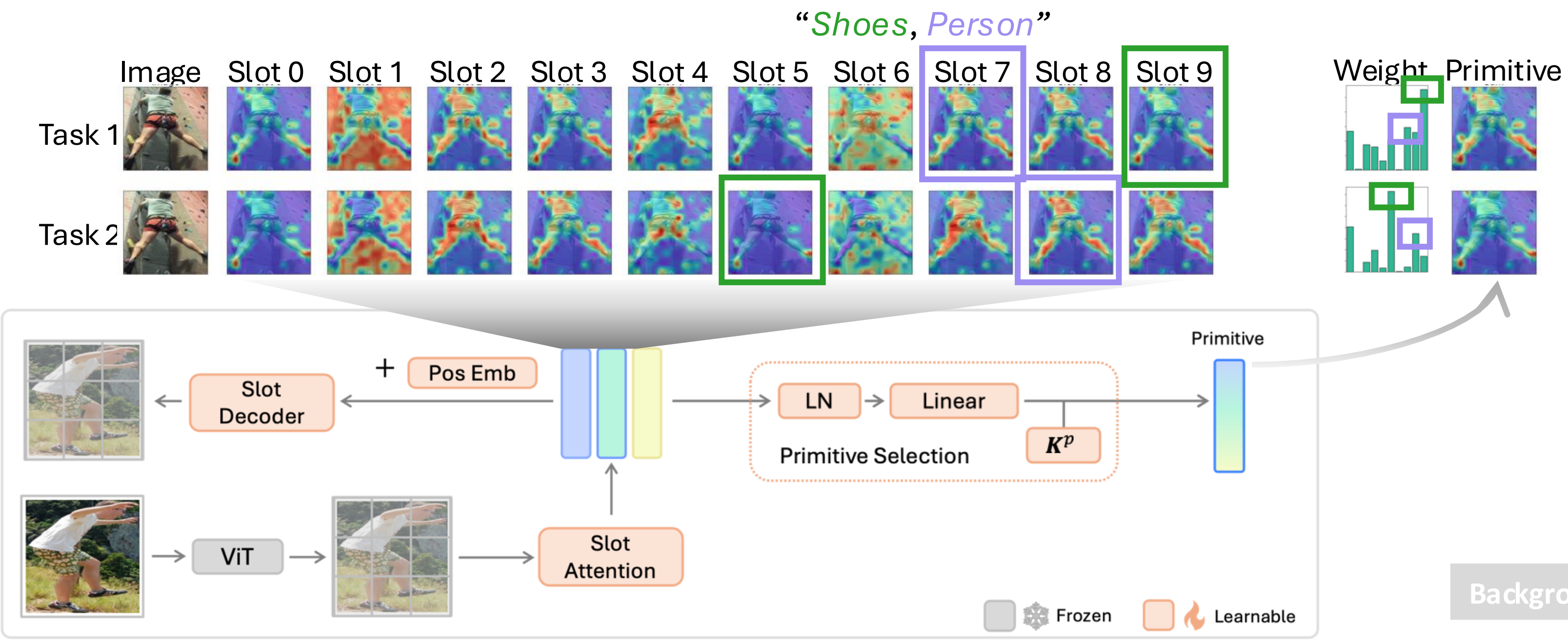
- Slot Attention: Decomposition into K independent slots
- Primitive Selection: Aggregating relevant slots into class-defining primitives



Stage 1: Concept Learning

Goal for challenge 1: Extracting conceptual image representations

- Slot Attention: Decomposition into K independent slots
- Primitive Selection: Aggregating relevant slots into class-defining primitives



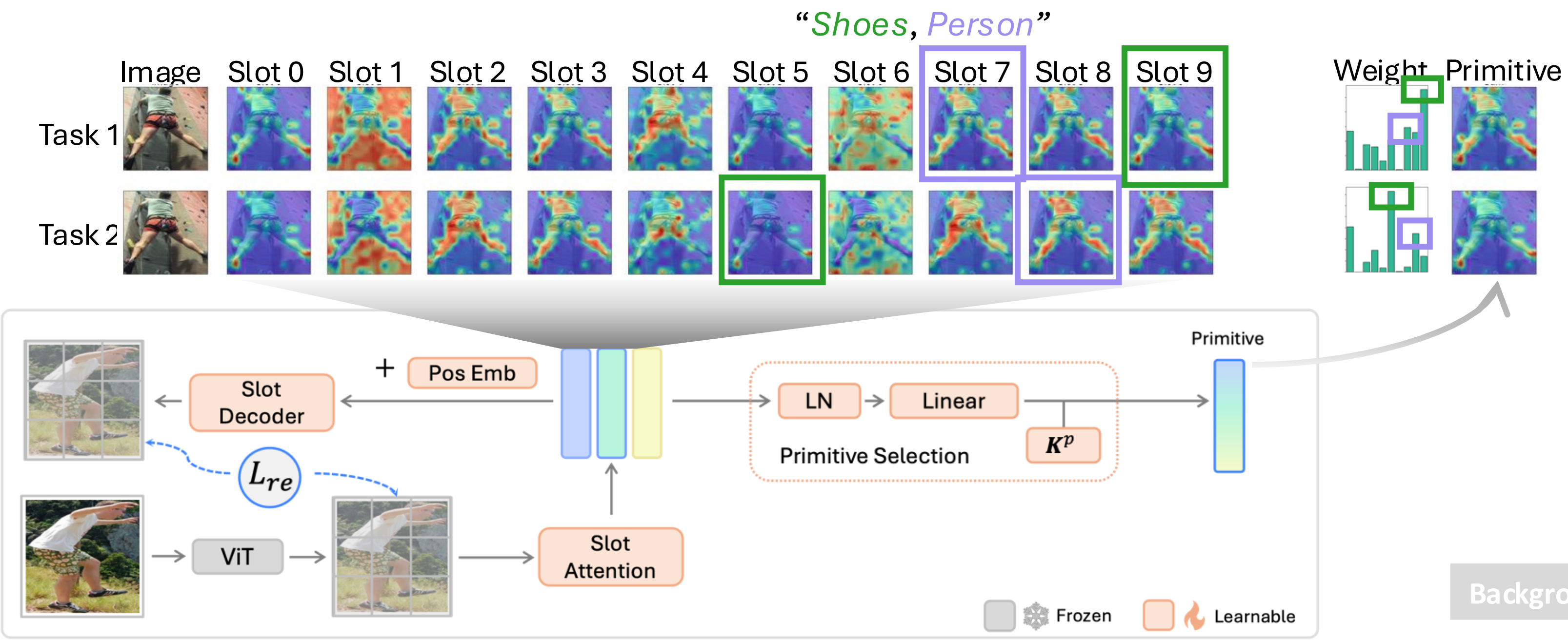
Stage 1: Concept Learning

Goal for challenge 1: Extracting conceptual image representations

- Slot Attention: Decomposition into K independent slots
- Primitive Selection: Aggregating relevant slots into class-defining primitives

Learning Objectives:

- Reconstruction Loss (L_{re}): Ensuring slots capture full semantic information



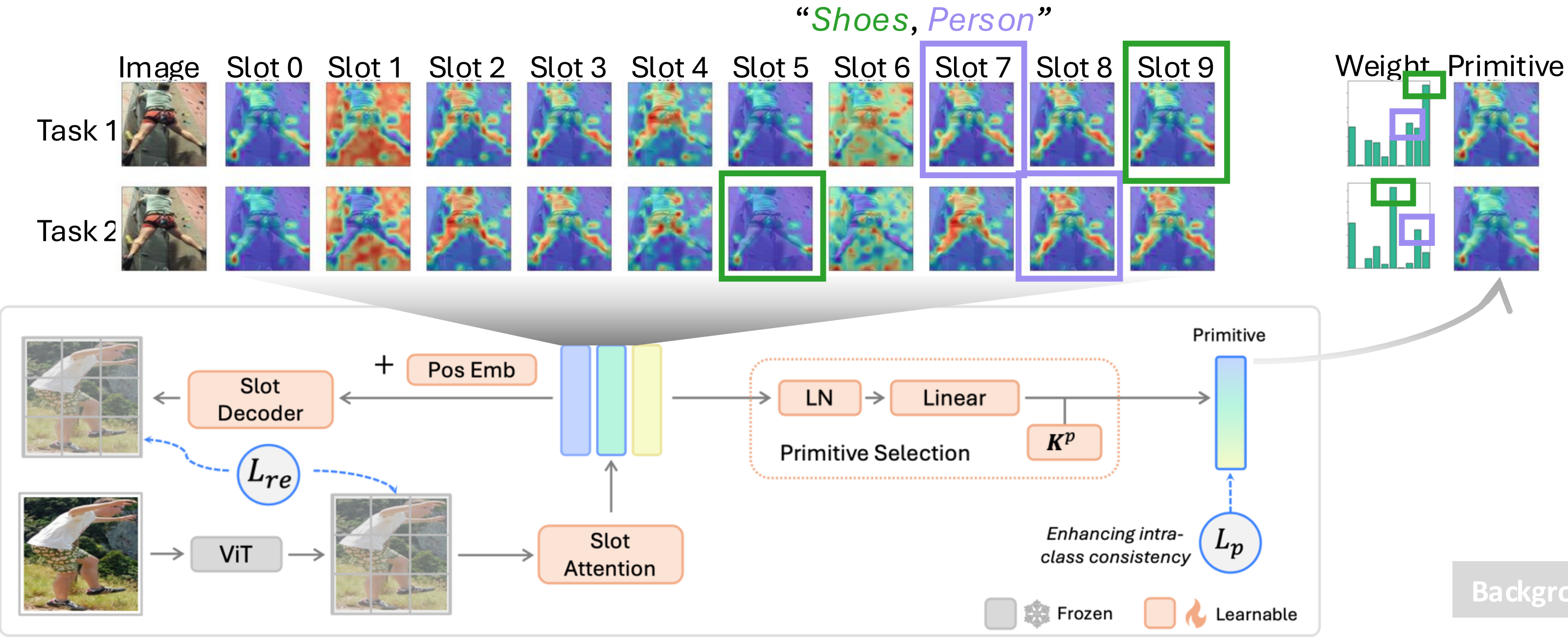
Stage 1: Concept Learning

Goal for challenge 1: Extracting conceptual image representations

- Slot Attention: Decomposition into K independent slots
- Primitive Selection: Aggregating relevant slots into class-defining primitives

Learning Objectives:

- Reconstruction Loss (L_{re}): Ensuring slots capture full semantic information
- Contrastive Primitive Loss (L_p): Minimizing intra-class primitive distance and maximizing inter-class primitive separation



Stage 1: Concept Learning

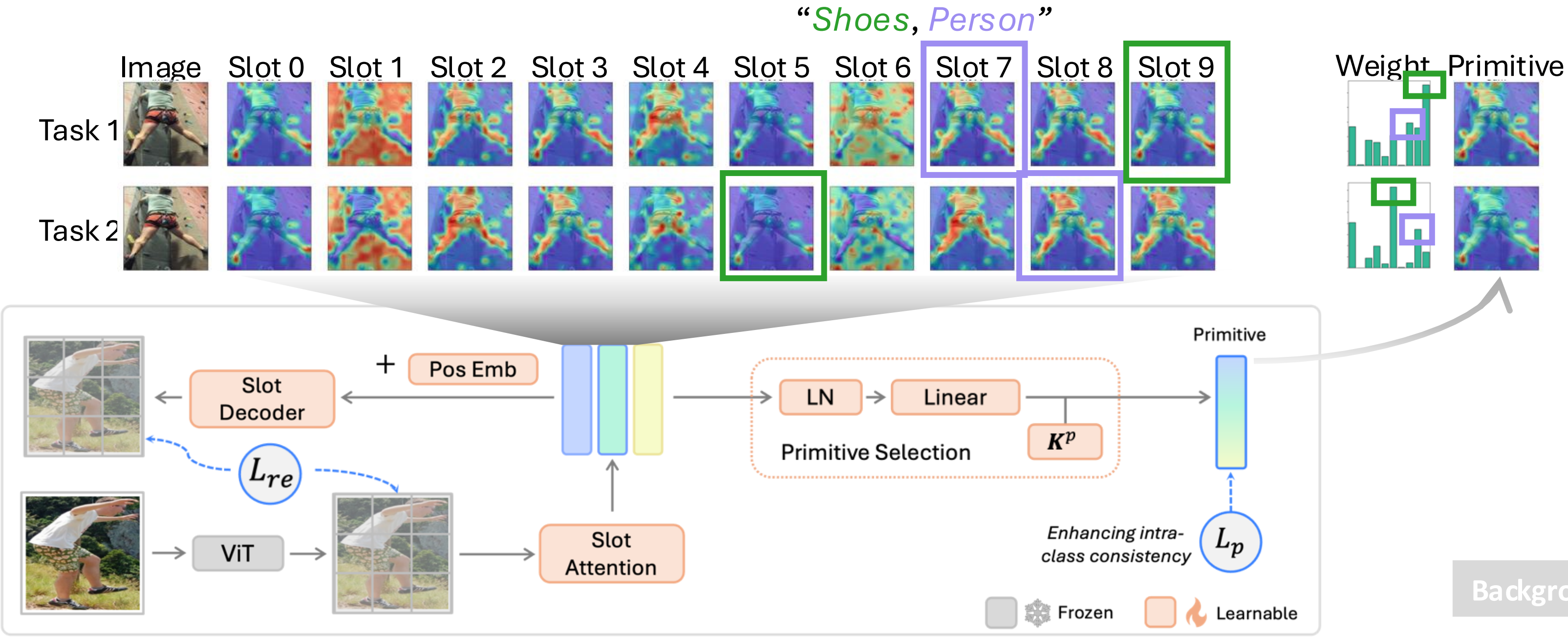
Goal for challenge 1: Extracting conceptual image representations

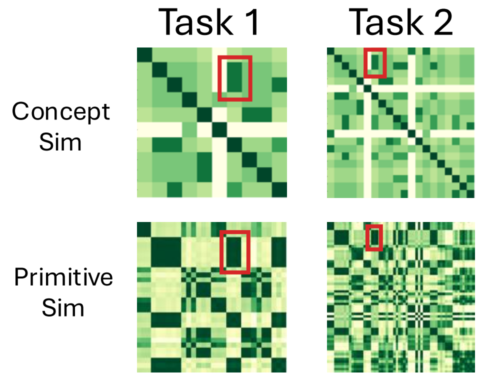
- Slot Attention: Decomposition into K independent slots
- Primitive Selection: Aggregating relevant slots into class-defining primitives

Learning Objectives:

- Reconstruction Loss (L_{re}): Ensuring slots capture full semantic information
- Contrastive Primitive Loss (L_p): Minimizing intra-class primitive distance and maximizing inter-class primitive separation

Outcome: Primitive mimics ground truth concept similarity without concept-level supervision.

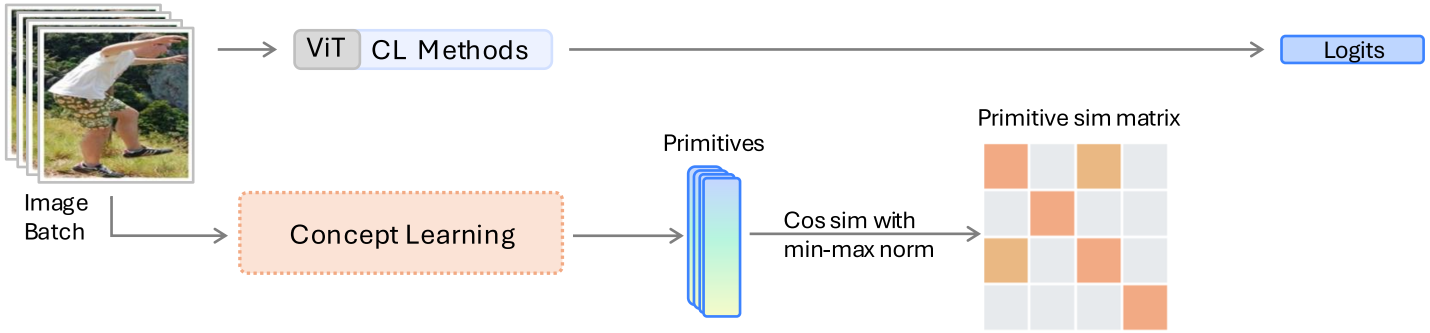




Class-wise similarity matrix

Stage 2: Method-Agnostic Primitive-Logit Knowledge Distillation

Goal for challenge 2: Grounding classifier decisions on structural concept compositions

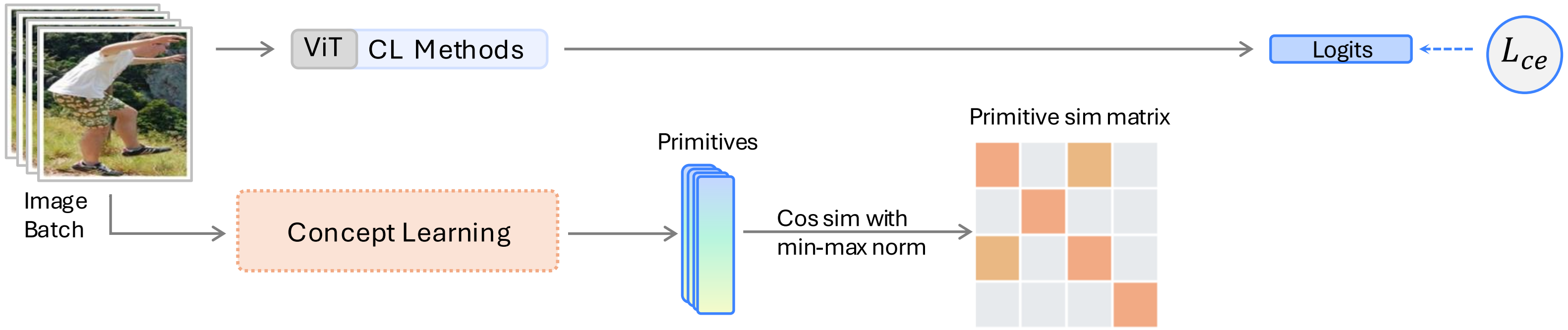


Stage 2: Method-Agnostic Primitive-Logit Knowledge Distillation

Goal for challenge 2: Grounding classifier decisions on structural concept compositions

Learning Objectives:

- CE Loss (L_{ce}): Maintaining task-specific discriminative power

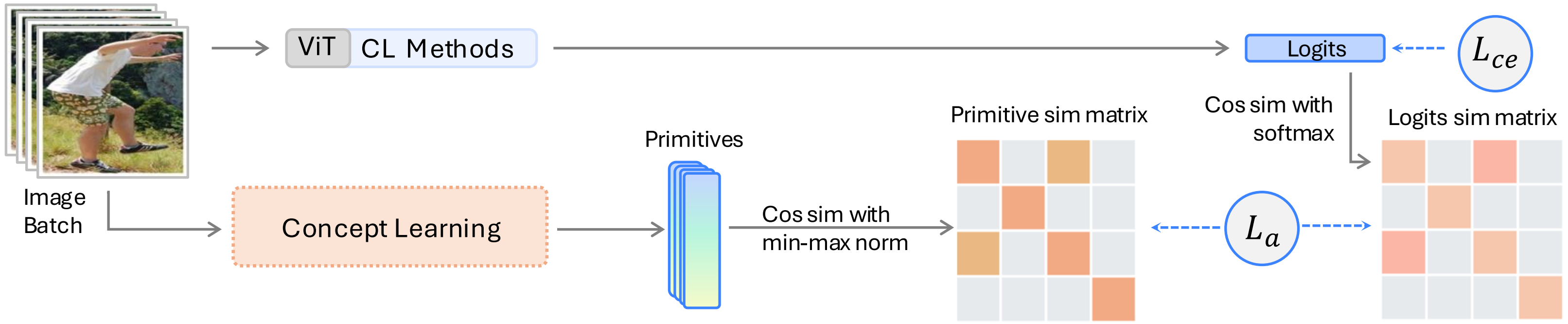


Stage 2: Method-Agnostic Primitive-Logit Knowledge Distillation

Goal for challenge 2: Grounding classifier decisions on structural concept compositions

Learning Objectives:

- CE Loss (L_{ce}): Maintaining task-specific discriminative power
- Contrastive Primitive-Logit Alignment Loss (L_a): Distilling sample-wise primitive similarities into logit distributions



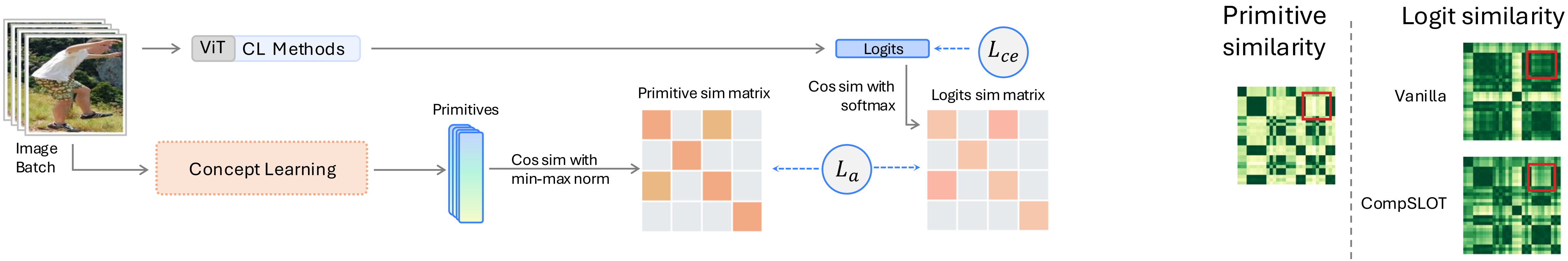
Stage 2: Method-Agnostic Primitive-Logit Knowledge Distillation

Goal for challenge 2: Grounding classifier decisions on structural concept compositions

Learning Objectives:

- CE Loss (L_{ce}): Maintaining task-specific discriminative power
- Contrastive Primitive-Logit Alignment Loss (L_a): Distilling sample-wise primitive similarities into logit distributions

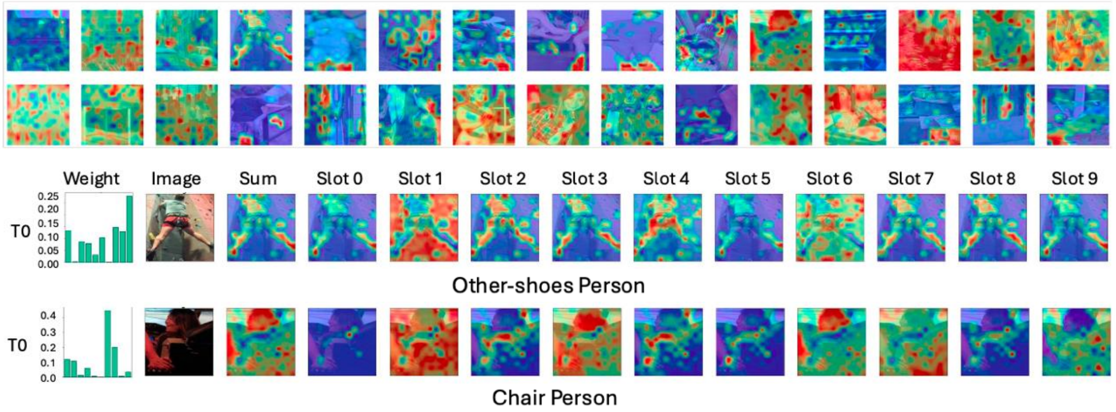
Outcome: CompSLOT distillates batch-wise primitive similarity, which carries concept relationship, into logits.



Challenge 1: Can CompSLOT extract latent concepts without external aid?

Challenge 1: Can CompSLOT extract latent concepts without external aid?

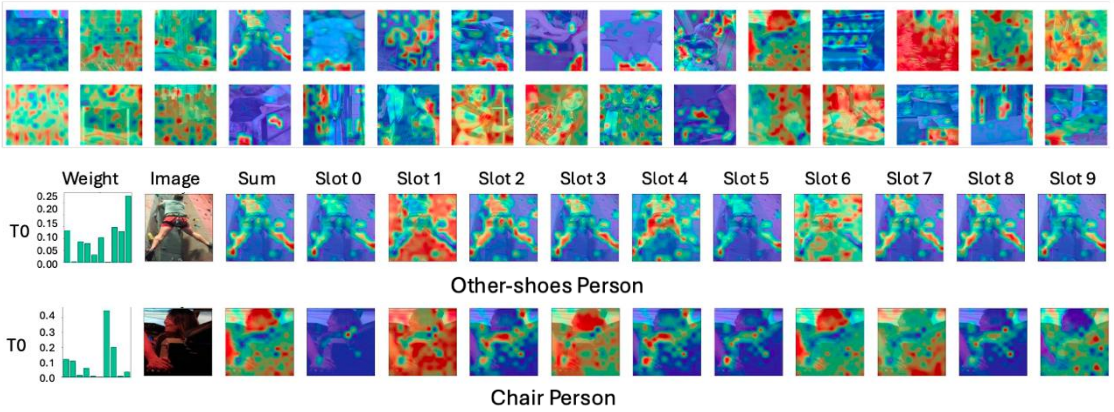
Takeaway 1: Slot Attention successfully decomposes images into semantically meaningful sub-parts.



Challenge 1: Can CompSLOT extract latent concepts without external aid?

Takeaway 1: Slot Attention successfully decomposes images into semantically meaningful sub-parts.

Takeaway 2: Primitive Selection successfully identifies important slots while filtering out background noise.

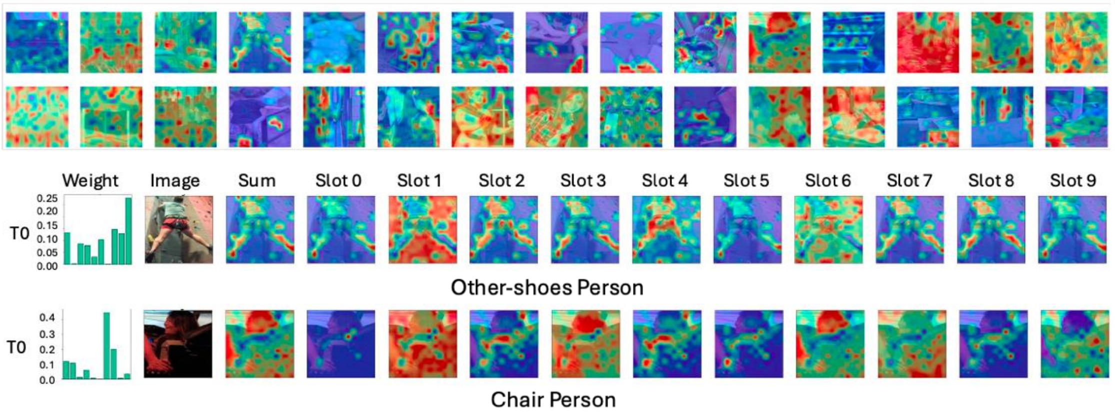


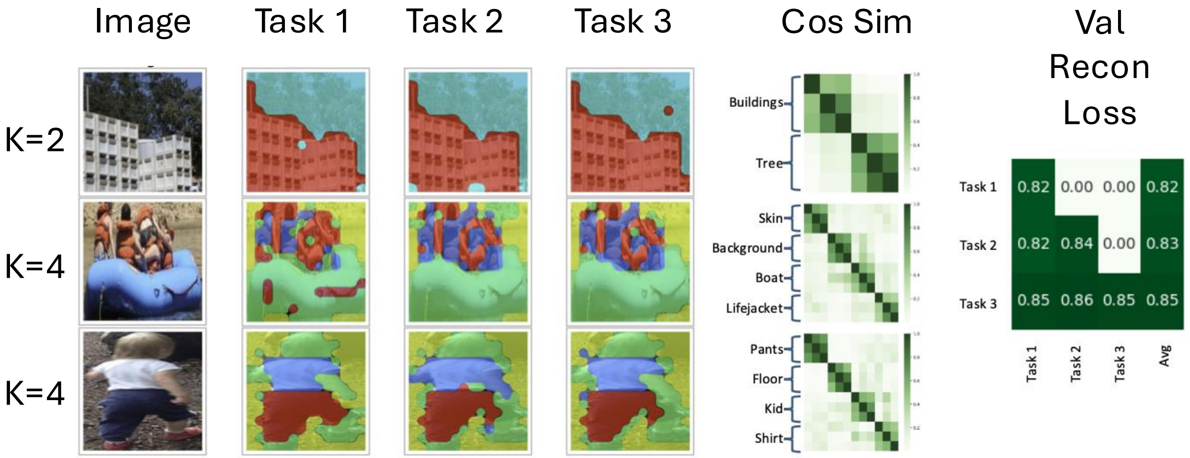
Challenge 1: Can CompSLOT extract latent concepts without external aid?

Takeaway 1: Slot Attention successfully decomposes images into semantically meaningful sub-parts.

Takeaway 2: Primitive Selection successfully identifies important slots while filtering out background noise.

Takeaway 3: Concept extraction remains stable after learning new tasks.





Challenge 2: Can CompSLOT benefit a broad range of CL algorithms?

Challenge 2: Can CompSLOT benefit a broad range of CL algorithms?

Takeaway 4: Consistent CL improvements (AA) and compositional generalization (CFST) across Prompt, Representation, and Mixture-based families.

Methods	Continual			CFST	
	AA (%) ↑	CA (%) ↑	FF (%) ↓	Hn (%) ↑	R↑
CPrompt	46.753±0.570	60.179±1.695	**15.670±0.950**	78.063±0.817	0.964
CPrompt †	**48.537±0.427**	**61.483±1.645**	18.315±1.111	**79.091±1.086**	**0.969**
ADAM + adapter	41.930±1.141	53.983±0.444	13.800±0.187	68.649±0.259	0.932
ADAM + adapter †	**49.480±1.201**	**60.989±0.641**	**12.896±0.379**	**74.335±0.572**	**0.958**
RanPAC	65.810±0.802	75.504±0.318	10.515±0.176	78.868±0.918	1.016
RanPAC †	**66.753±0.867**	**76.584±0.603**	**10.219±0.281**	**79.815±0.829**	**1.032**
EASE	47.657±1.494	59.475±2.574	**18.215±0.107**	79.713±0.449	0.996
EASE †	**49.323±1.165**	**62.603±1.252**	22.470±2.472	**82.887±0.320**	**1.001**
CoFiMA	65.107±0.508	73.227±1.047	15.248±0.542	86.711±0.483	1.011
CoFiMA †	**66.170±0.578**	**74.322±0.463**	**14.204±0.880**	**88.297±0.278**	**1.017**
FOSTER*	60.863±0.271	68.800±0.496	**2.441±0.122**	89.791±0.086	1.087
FOSTER* †	**66.290±1.451**	**71.828±2.619**	6.470±5.770	**89.910±0.710**	**1.154**
DER*	52.003±1.019	62.675±1.695	40.122±0.907	**90.119±0.510**	1.080
DER* †	**54.900±1.093**	**66.020±1.049**	**38.941±0.995**	88.986±0.129	**1.096**
MEMO*	56.553±1.804	66.462±0.702	9.289±0.326	82.425±1.282	1.029
MEMO* †	**58.653±1.449**	**68.037±1.459**	**8.944±0.268**	**84.003±1.451**	**1.050**

Challenge 2: Can CompSLOT benefit a broad range of CL algorithms?

Takeaway 4: Consistent CL improvements (AA) and compositional generalization (CFST) across Prompt, Representation, and Mixture-based families.

Takeaway 5: The performance gains are not the outcome of increased model capacity.

Methods	AA (%) ↑	Continual CA (%) ↑	FF (%) ↓	CFST Hn (%) ↑	R↑
CPrompt	46.753±0.570	60.179±1.695	**15.670±0.950**	78.063±0.817	0.964
CPrompt †	**48.537±0.427**	**61.483±1.645**	18.315±1.111	**79.091±1.086**	**0.969**
ADAM + adapter	41.930±1.141	53.983±0.444	13.800±0.187	68.649±0.259	0.932
ADAM + adapter †	**49.480±1.201**	**60.989±0.641**	**12.896±0.379**	**74.335±0.572**	**0.958**
RanPAC	65.810±0.802	75.504±0.318	10.515±0.176	78.868±0.918	1.016
RanPAC †	**66.753±0.867**	**76.584±0.603**	**10.219±0.281**	**79.815±0.829**	**1.032**
EASE	47.657±1.494	59.475±2.574	**18.215±0.107**	79.713±0.449	0.996
EASE †	**49.323±1.165**	**62.603±1.252**	22.470±2.472	**82.887±0.320**	**1.001**
CoFiMA	65.107±0.508	73.227±1.047	15.248±0.542	86.711±0.483	1.011
CoFiMA †	**66.170±0.578**	**74.322±0.463**	**14.204±0.880**	**88.297±0.278**	**1.017**
FOSTER*	60.863±0.271	68.800±0.496	**2.441±0.122**	89.791±0.086	1.087
FOSTER* †	**66.290±1.451**	**71.828±2.619**	6.470±5.770	**89.910±0.710**	**1.154**
DER*	52.003±1.019	62.675±1.695	40.122±0.907	**90.119±0.510**	1.080
DER* †	**54.900±1.093**	**66.020±1.049**	**38.941±0.995**	88.986±0.129	**1.096**
MEMO*	56.553±1.804	66.462±0.702	9.289±0.326	82.425±1.282	1.029
MEMO* †	**58.653±1.449**	**68.037±1.459**	**8.944±0.268**	**84.003±1.451**	**1.050**

Methods	L_p	L_a	AA (%) ↑	R↑
RanPAC	✗+param	✗	65.080	1.010
	✗avg	✓	58.220	0.969
	✓avg	✓	65.870	1.003
	✓sig	✓	65.950	1.020
	✓sign	✓	65.140	1.006
	✓cos	✓	63.910	0.989
	✓soft	✓	**66.753**	**1.032**
CPrompt	✗+param	✗	46.300	**0.969**
	✗avg	✓	40.230	0.952
	✓avg	✓	47.690	0.958
	✓sig	✓	48.080	0.961
	✓sign	✓	47.780	0.966
	✓cos	✓	47.410	0.964
	✓soft	✓	**48.537**	**0.969**

Challenge 2: Can CompSLOT benefit a broad range of CL algorithms?

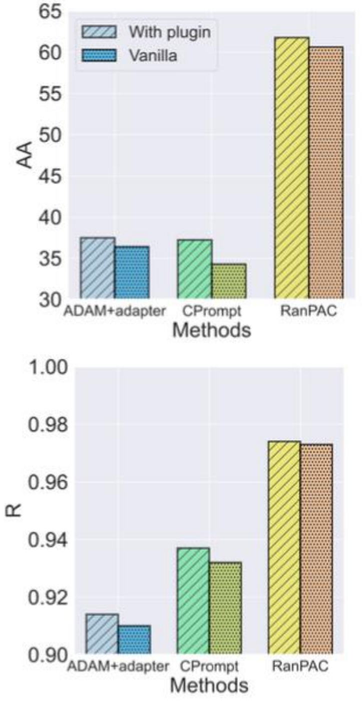
Takeaway 4: Consistent CL improvements (AA) and compositional generalization (CFST) across Prompt, Representation, and Mixture-based families.

Takeaway 5: The performance gains are not the outcome of increased model capacity.

Takeaway 6: Long-sequence robustness over challenging 20 tasks.

Methods	AA (%) ↑	Continual CA (%) ↑	FF (%) ↓	CFST Hn (%) ↑	R↑
CPrompt	46.753±0.570	60.179±1.695	**15.670±0.950**	78.063±0.817	0.964
CPrompt †	**48.537±0.427**	**61.483±1.645**	18.315±1.111	**79.091±1.086**	**0.969**
ADAM + adapter	41.930±1.141	53.983±0.444	13.800±0.187	68.649±0.259	0.932
ADAM + adapter †	**49.480±1.201**	**60.989±0.641**	**12.896±0.379**	**74.335±0.572**	**0.958**
RanPAC	65.810±0.802	75.504±0.318	10.515±0.176	78.868±0.918	1.016
RanPAC †	**66.753±0.867**	**76.584±0.603**	**10.219±0.281**	**79.815±0.829**	**1.032**
EASE	47.657±1.494	59.475±2.574	**18.215±0.107**	79.713±0.449	0.996
EASE †	**49.323±1.165**	**62.603±1.252**	22.470±2.472	**82.887±0.320**	**1.001**
CoFiMA	65.107±0.508	73.227±1.047	15.248±0.542	86.711±0.483	1.011
CoFiMA †	**66.170±0.578**	**74.322±0.463**	**14.204±0.880**	**88.297±0.278**	**1.017**
FOSTER*	60.863±0.271	68.800±0.496	**2.441±0.122**	89.791±0.086	1.087
FOSTER* †	**66.290±1.451**	**71.828±2.619**	6.470±5.770	**89.910±0.710**	**1.154**
DER*	52.003±1.019	62.675±1.695	40.122±0.907	**90.119±0.510**	1.080
DER* †	**54.900±1.093**	**66.020±1.049**	**38.941±0.995**	88.986±0.129	**1.096**
MEMO*	56.553±1.804	66.462±0.702	9.289±0.326	82.425±1.282	1.029
MEMO* †	**58.653±1.449**	**68.037±1.459**	**8.944±0.268**	**84.003±1.451**	**1.050**

Methods	L_p	L_a	AA (%) ↑	R↑
RanPAC	✗+param	✗	65.080	1.010
	✗avg	✓	58.220	0.969
	✓avg	✓	65.870	1.003
	✓sig	✓	65.950	1.020
	✓sign	✓	65.140	1.006
	✓cos	✓	63.910	0.989
	✓soft	✓	**66.753**	**1.032**
CPrompt	✗+param	✗	46.300	**0.969**
	✗avg	✓	40.230	0.952
	✓avg	✓	47.690	0.958
	✓sig	✓	48.080	0.961
	✓sign	✓	47.780	0.966
	✓cos	✓	47.410	0.964
	✓soft	✓	**48.537**	**0.969**



Takeaways

High-dimensional un-structural feature space causes neglect of essential concepts for easy comparison paths.

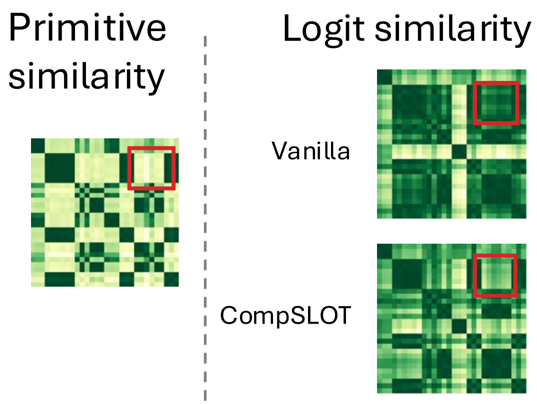
- Solution: Grounding decisions on structural, low-dim concept space

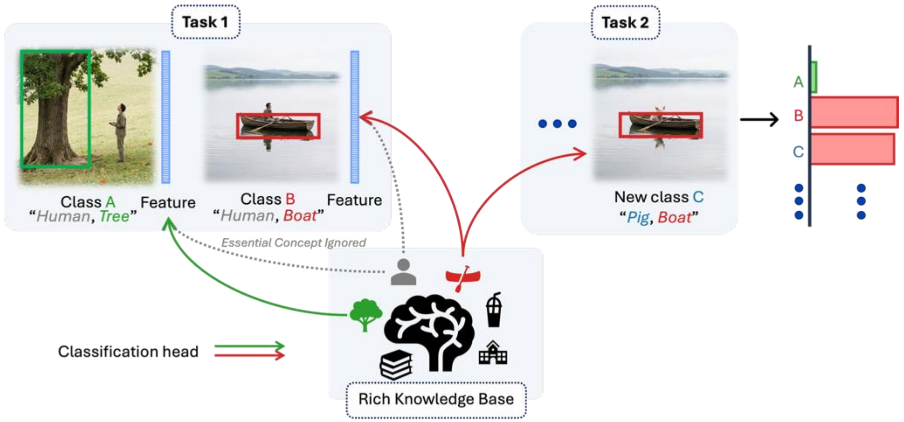
Challenge 1: Is there inherent concept awareness in Foundation Model?

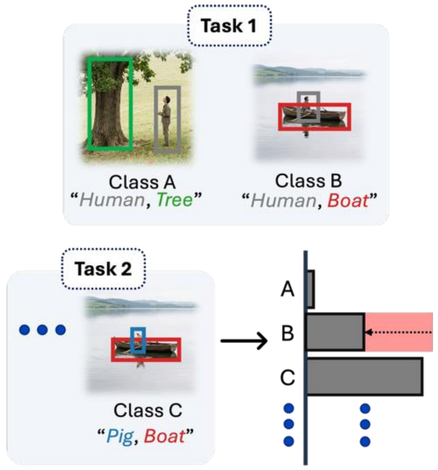
- CompSLOT uses Slot Attention and primitive selection.

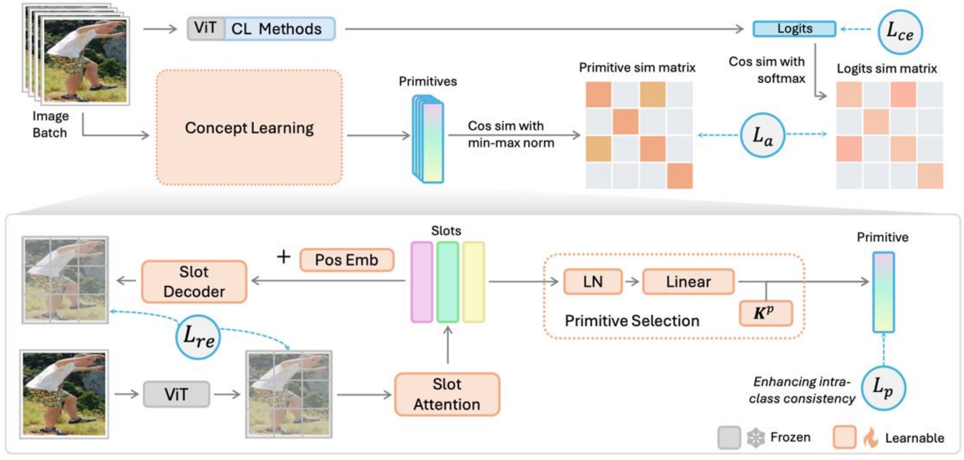
Challenge 2: How to establish cross-class concept grounding for a wide range of CL algorithms?

- CompSLOT distillates pair-wise primitive similarity into logits.

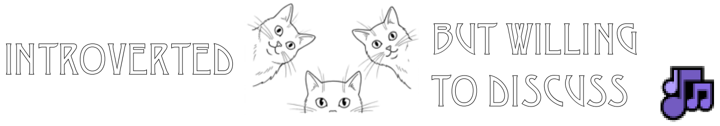








Poster: **Pavilion 3 P3-#304**,
on **Sat, Apr 25, 2026 • 3:15 PM – 5:45 PM -03.**



https://github.com/liaoweiduo/CompSLOT