



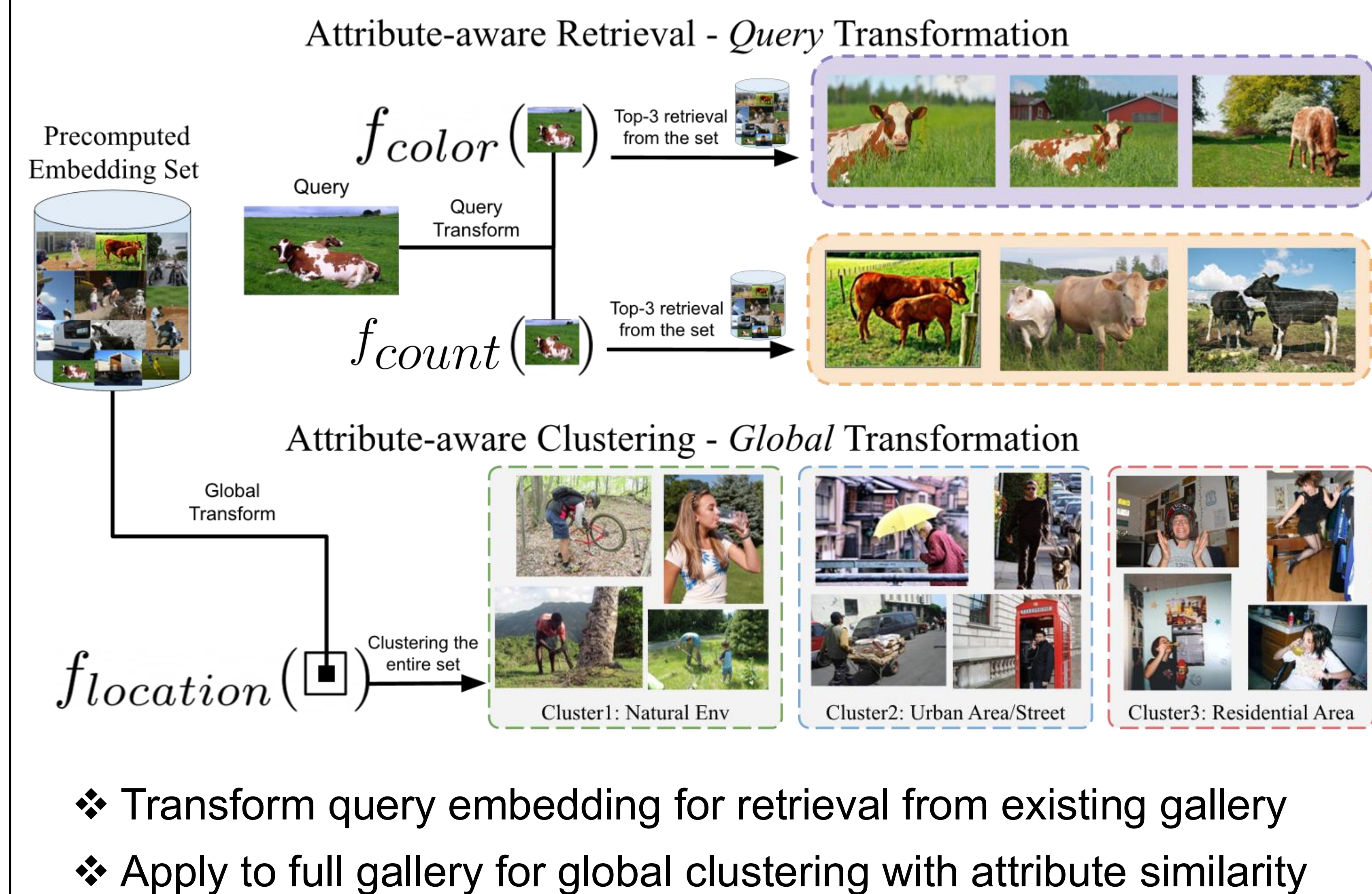
Motivation – Attribute-Aware Embedding Space

- ❖ Embeddings primarily express dominant semantic information
- ❖ They contain but suppress other important attributes like color

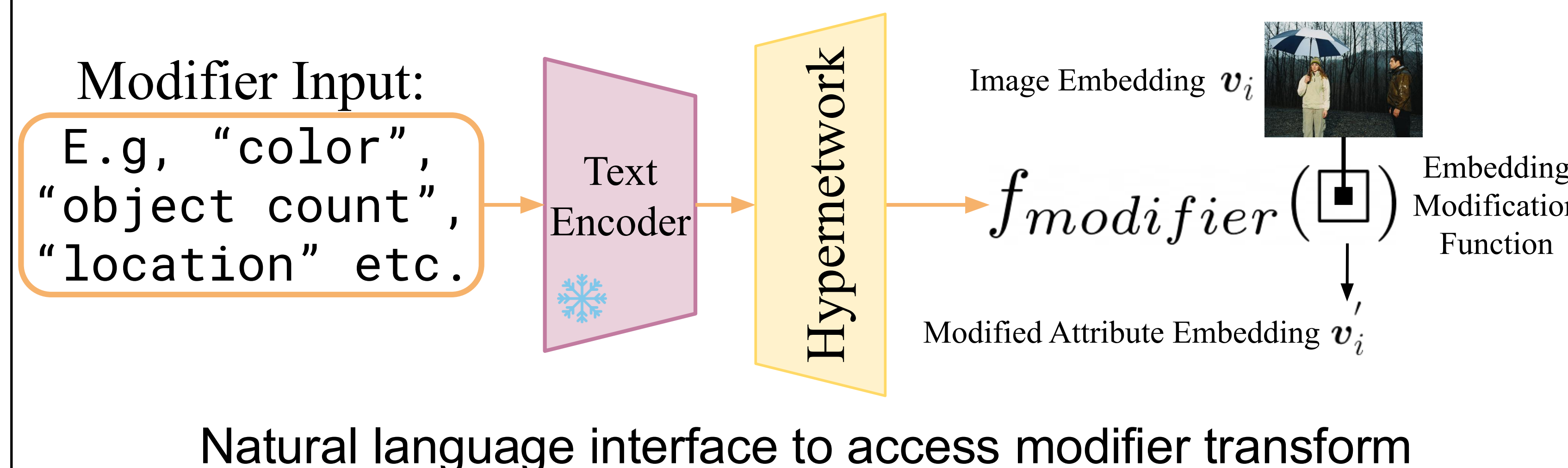
Contributions

- ❖ Learn transformations to emphasize each individual attribute
- ❖ Capable of customizable retrieval from existing galleries
- ❖ Proposed transform induces attribute-specific similarity metric
- ❖ Transforms generated from trained lightweight hypernetwork

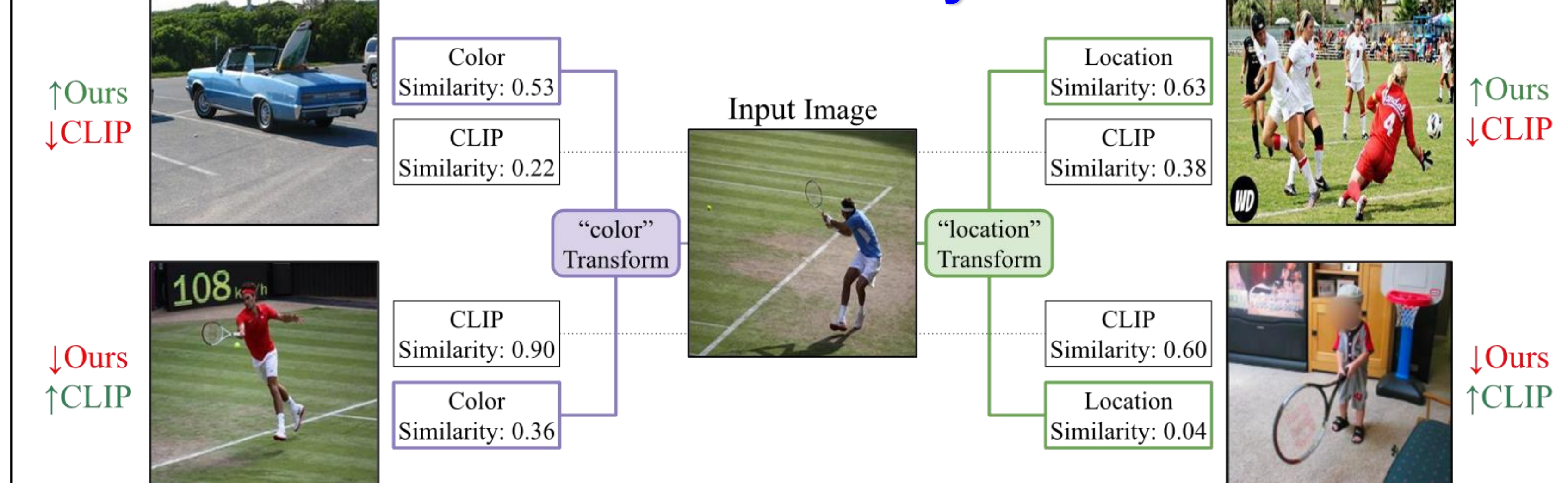
Multi-Use Embedding Transformation Settings



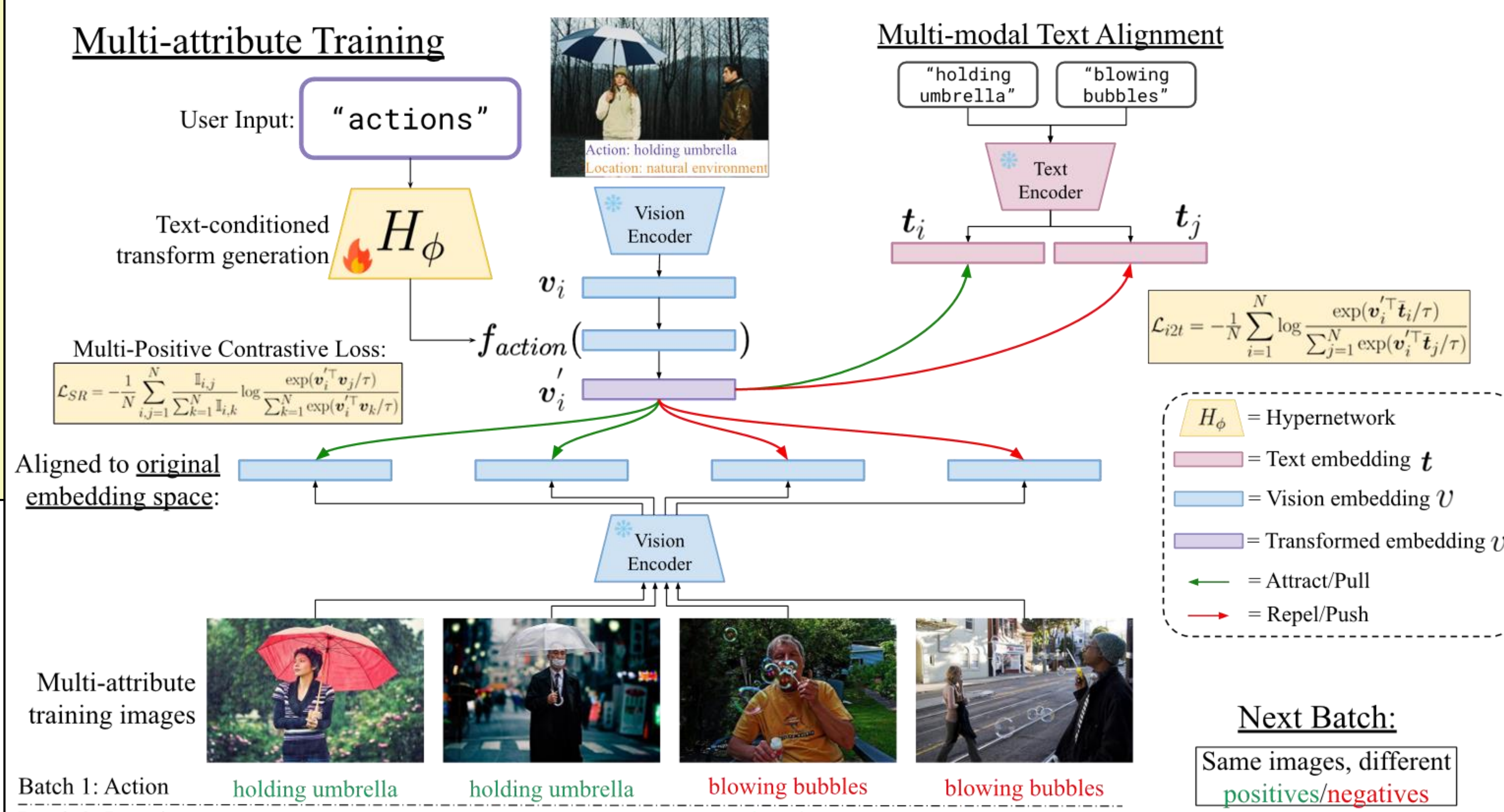
Generating Transforms



Attribute Similarity Score



Training Workflow Diagram

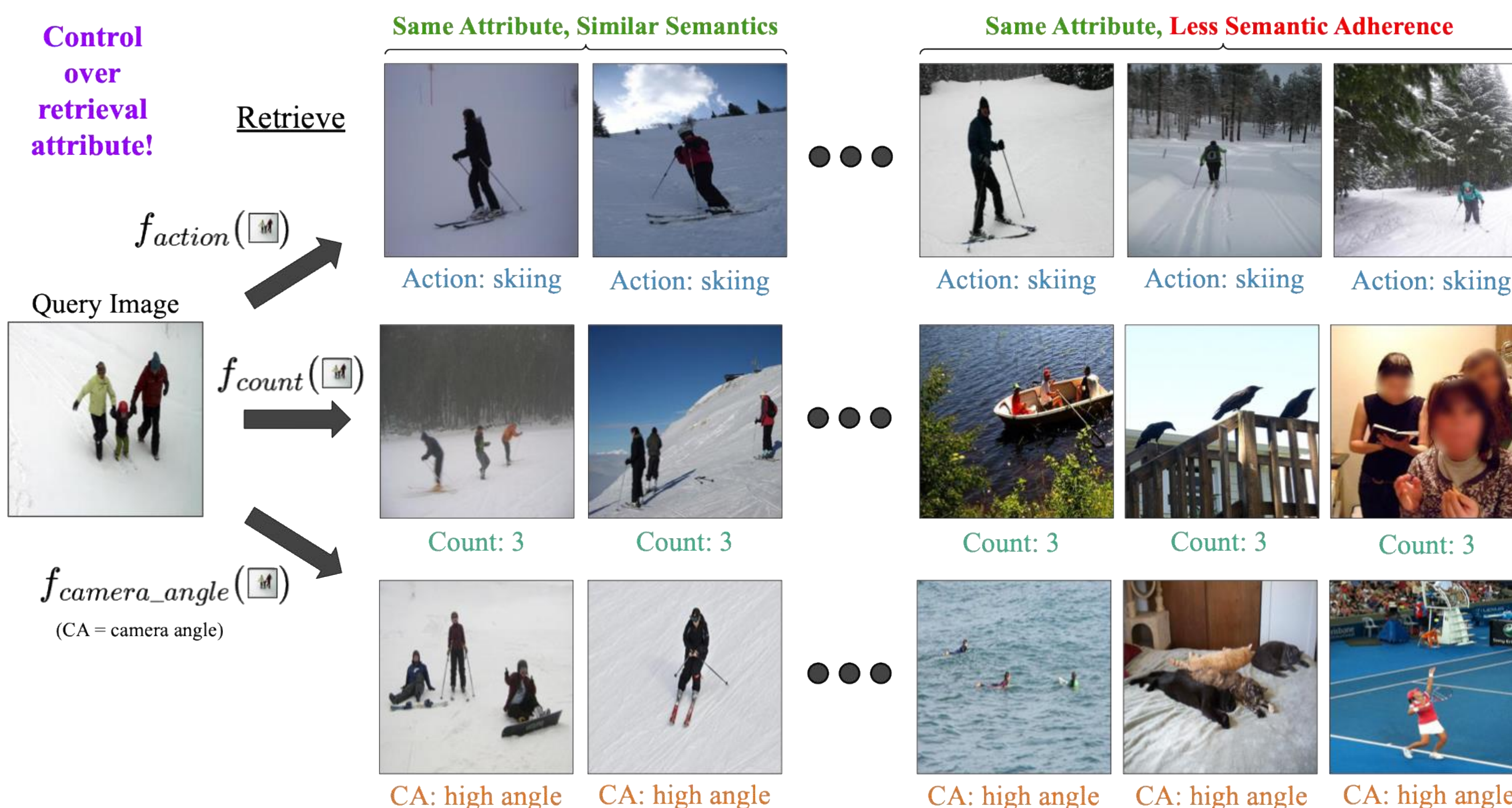


Results

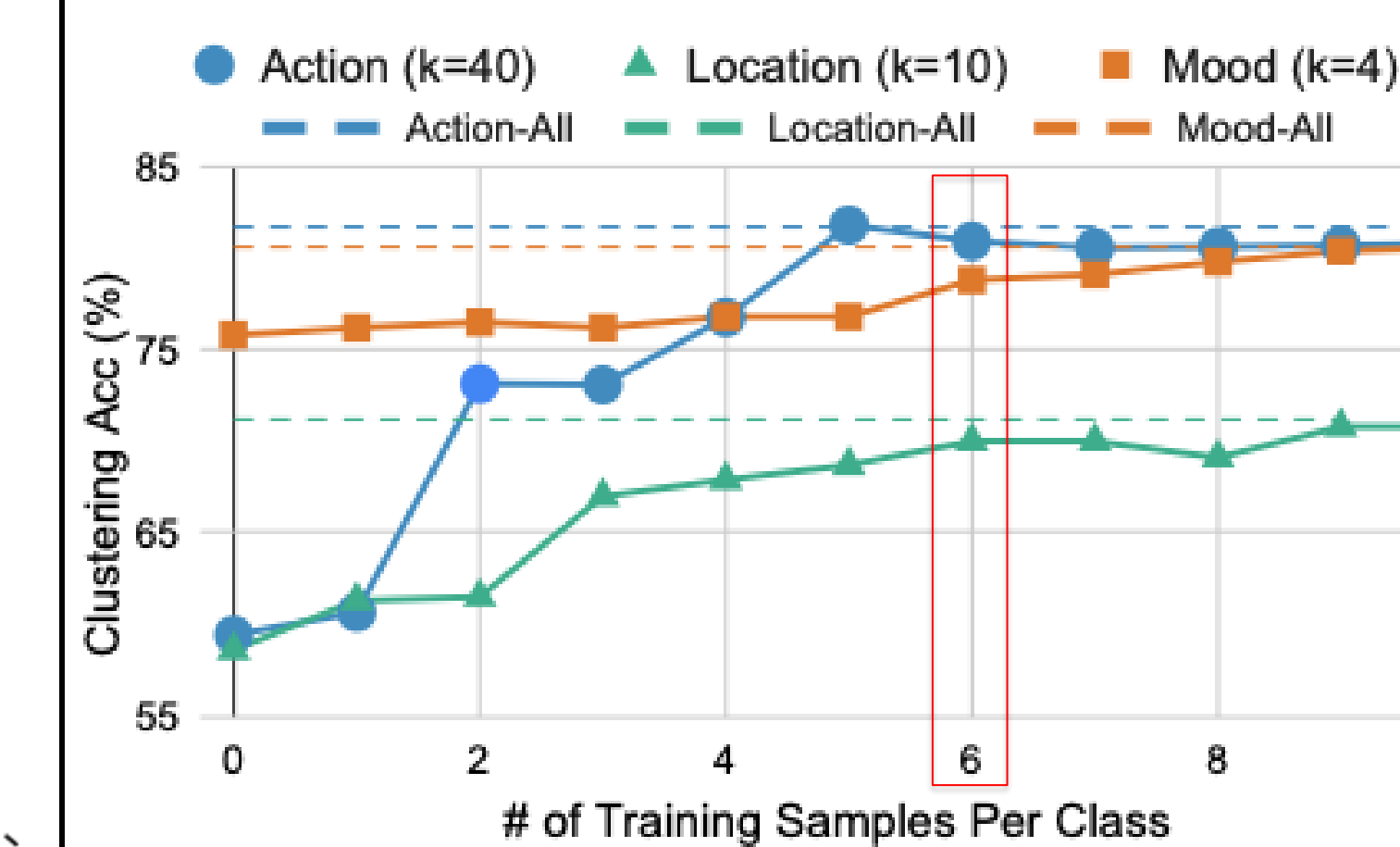
Attribute-aware Retrieval (frozen gallery)					Attribute Clustering				
Method	Multi-Attr. Capability	Clevr4 mAP	S40 mAP	ShotBench mAP	Method	Inference Time (s)	Clevr4 c. ACC	S40 c. ACC	ShotBench c. ACC
Base Features	✗	33.8	43.7	25.2	Base Features	0.59	48.8	65.2	20.8
LCR	✓	15.9	19.9	18.9	LCR	0.62	56.4	64.8	30.2
CWP	✓	43.8	49.8	21.4	CWP	0.68	60.3	64.9	32.2
ArcFace	✗	12.3	18.8	17.5	Multi-MAP	750	62.8	62.8	32.6
SupCon	✗	12.7	18.8	17.7	Multi-Sub	15300	72.2	66.1	32.4
StableRep+	✗	49.5	21.3	20.9	IC TC	37680	57.9	76.1	21.7
MaPLe	✗	33.5	47.8	23.7	SSD-LLM	27600	56.5	74.1	25.0
LoRA	✗	42.9	49.3	22.0	X-Cluster	104760	64.9	68.3*	-
SEARLE-XL	✗	24.6	33.9	20.3	SEARLE-XL	81	30.7	46.9	24.5
Ours-Query	✓	77.0	86.3	41.0	Ours-Space	0.64	73.0	80.0	51.6

3 datasets, 14 attribute categories, 187 subclasses

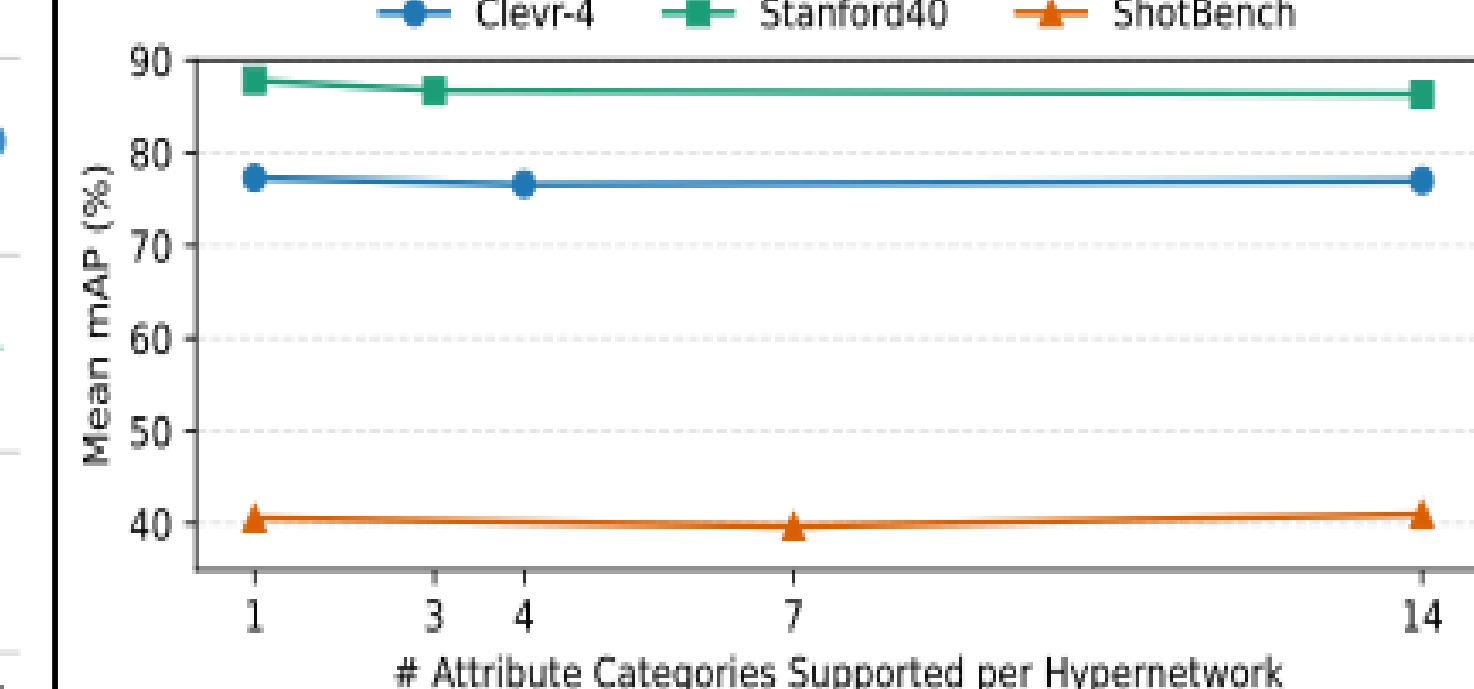
Qualitative Retrieval Examples



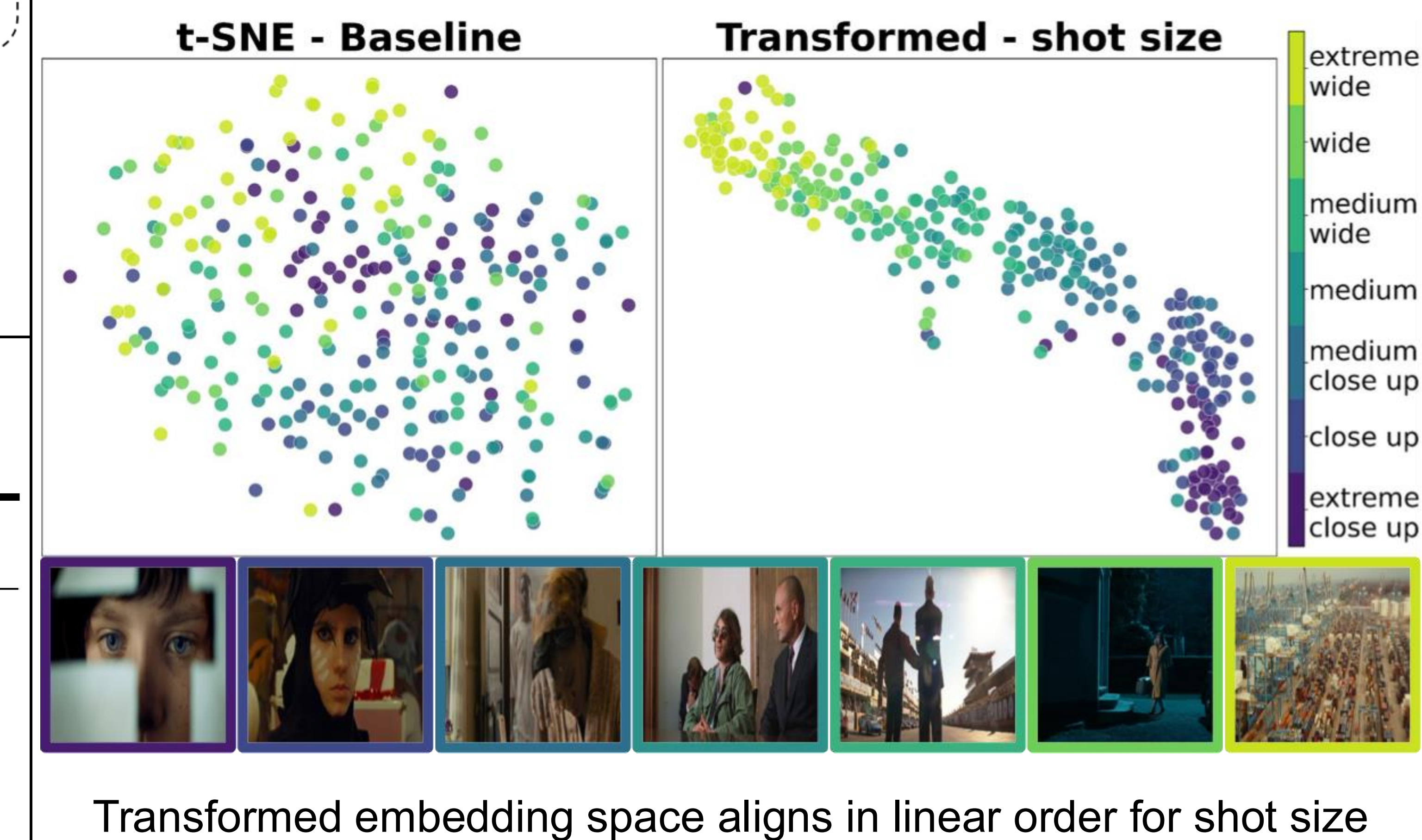
Low Data Requirement



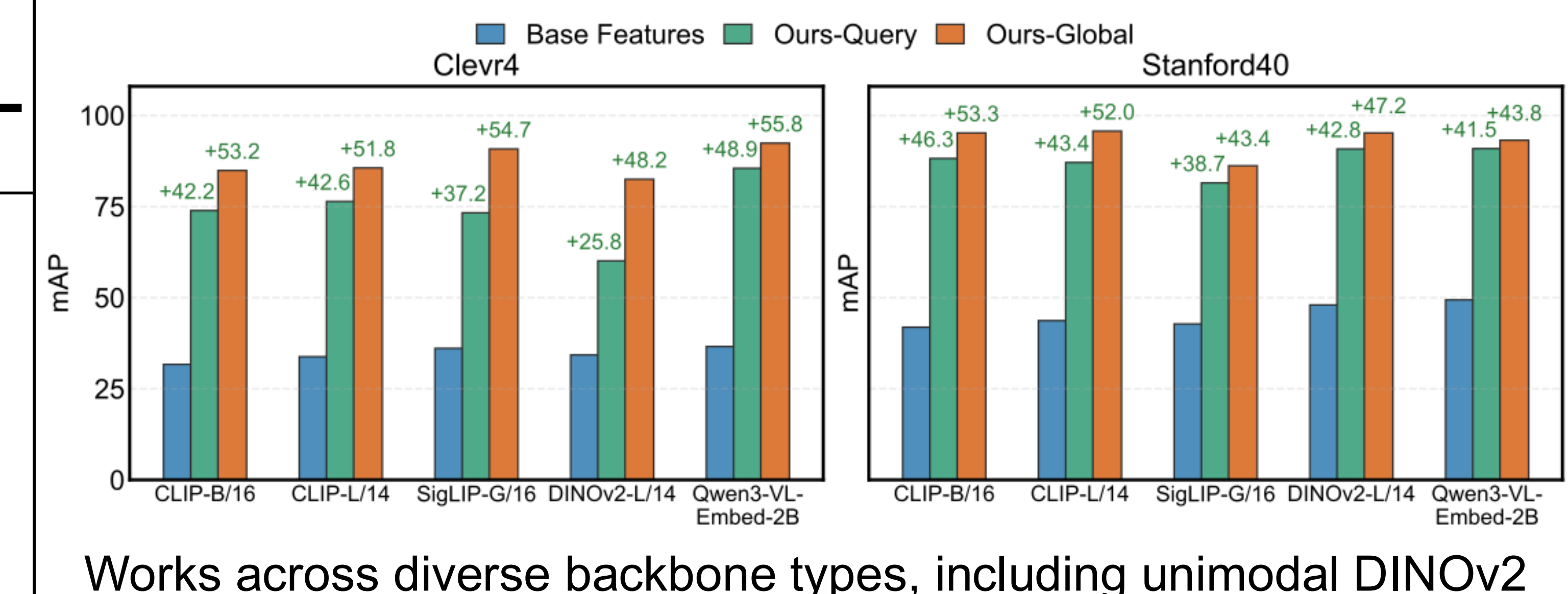
Hypernetwork Capacity



t-SNE Visualization



Backbone Generalization



Compositional Retrieval

Blended query: $q_{blend} = \alpha (W_a q + b_a) + (1 - \alpha) (W_b q + b_b)$

Variant	Neither@5	L-only@5	A-only@5	Both@5	mAP (Both)
Base	17.5	15.9	21.5	45.1	48.5
Location only	5.3	76.4	2.8	15.6	21.2
Action only	1.0	0.9	40.7	57.4	70.0
Linear Blend ($\alpha = 0.5$)	1.4	3.6	30.2	64.8	80.9

Uniform blending improves joint-label retrieval without retraining