

Triple Correlations

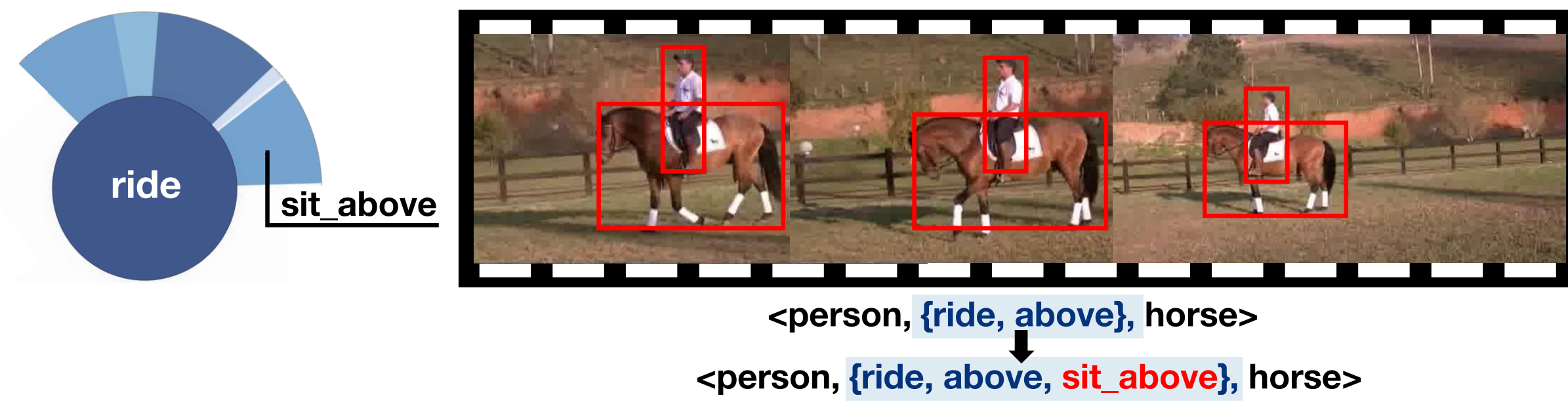
Motivation:

Due to the inherently biased distribution and missing annotations in the training data, current VidSGG methods have been found to perform poorly on less-represented predicates.

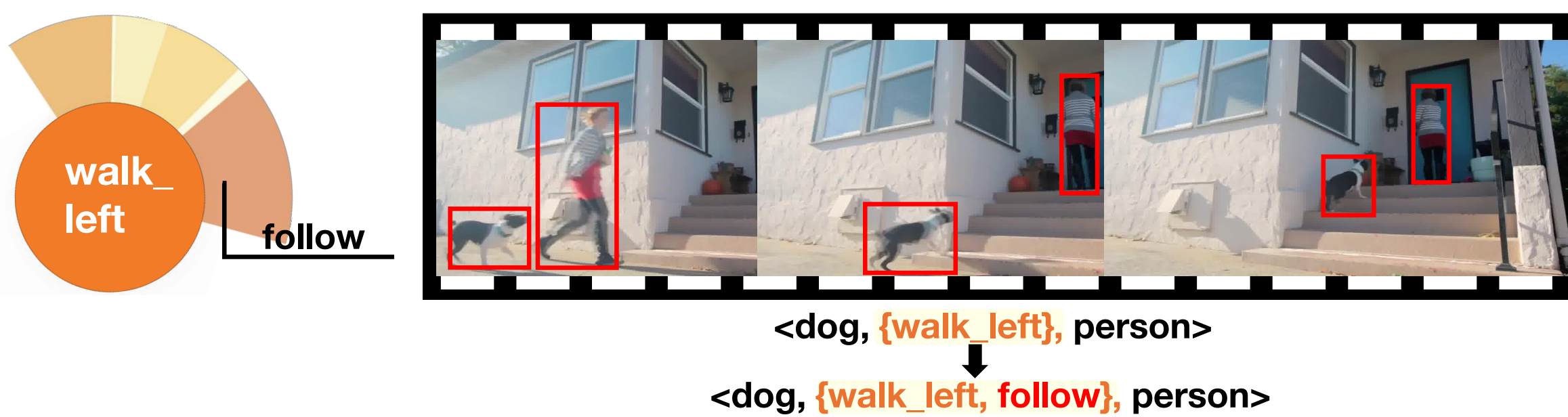
Contributions:

- the first method to address VidSGG from an explicit perspective of missing label supplementation.
- use triple complementary correlations to guide the label supplementation process, ensuring that the missing labels are effectively supplemented to achieve an unbiased graph generation.

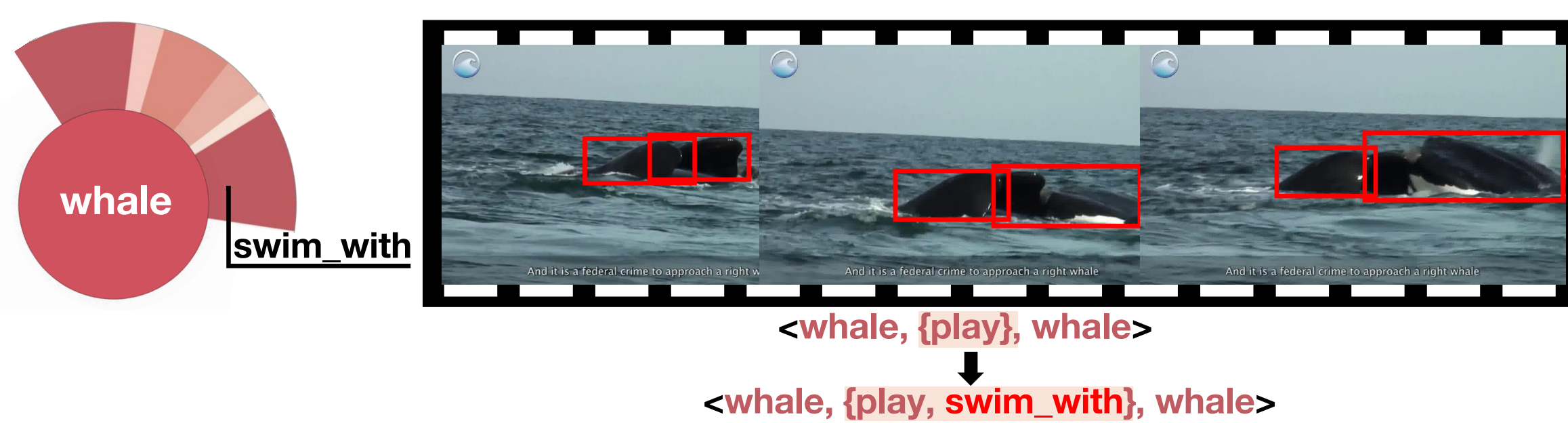
(a) Spatial predicates-predicates correlations



(b) Temporal predicates-predicates correlations

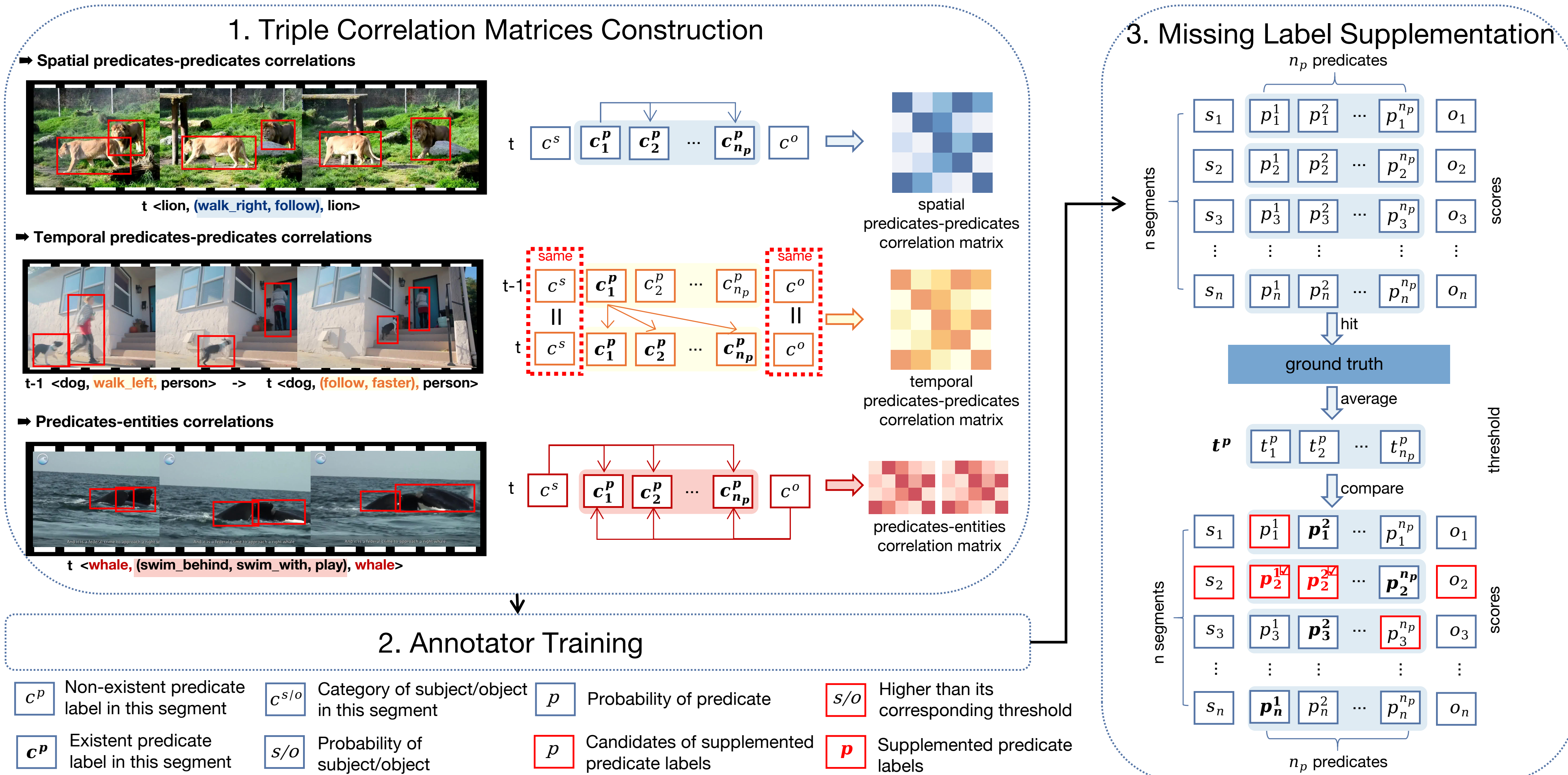


(c) Predicates-entities correlations



Triple Correlations-Guided Label Supplementation

The predicate labels that are supplemented are directly applied to the training set.



Experiments Results

Performance of Trico on VidVRD dataset in PredCls.

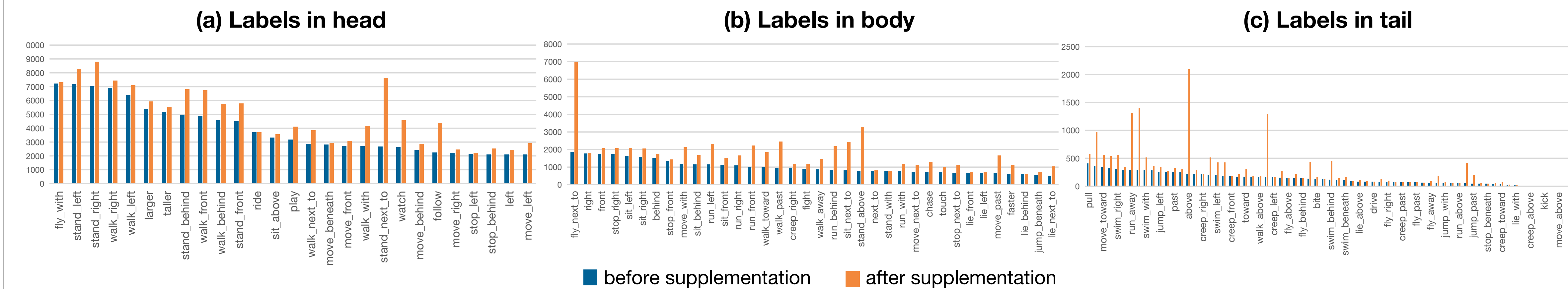
Method	Relation Detection						Relation Tagging	
	mR@50	mR@100	R@50	R@100	Mean	mAP	P@5	P@10
VidVRD-II [30] [*]	37.09	45.45	44.43	59.28	46.56	47.32	47.30	36.50
+Trico (ours)	36.57	48.10	44.63	59.65	47.24	48.47	48.20	36.40
+Trico+LoS (ours)	37.21	49.15	43.35	59.38	47.27	48.17	48.60	37.45

^{*} We re-implemented VidVRD-II and fixed the fake performance boost issue of the original evaluation API.

Ablation on VidVRD dataset in PredCls.

Method	Head		Body		Tail
	mR@100	mR@100	mR@100	mR@100	mAP
VidVRD-II [30]	45.35	78.08	49.55	30.36	47.49
+base	45.84	78.83	54.82	28.24	46.15
+T (ours)	47.66	78.33	51.52	33.65	46.43
+T+S (ours)	48.04	74.74	53.60	34.70	47.40
+T+S+E (ours)	48.10	76.46	56.90	32.77	48.47

Statistical Results of Label Supplementation



Visualization of the VidSGG on VidVRD dataset

