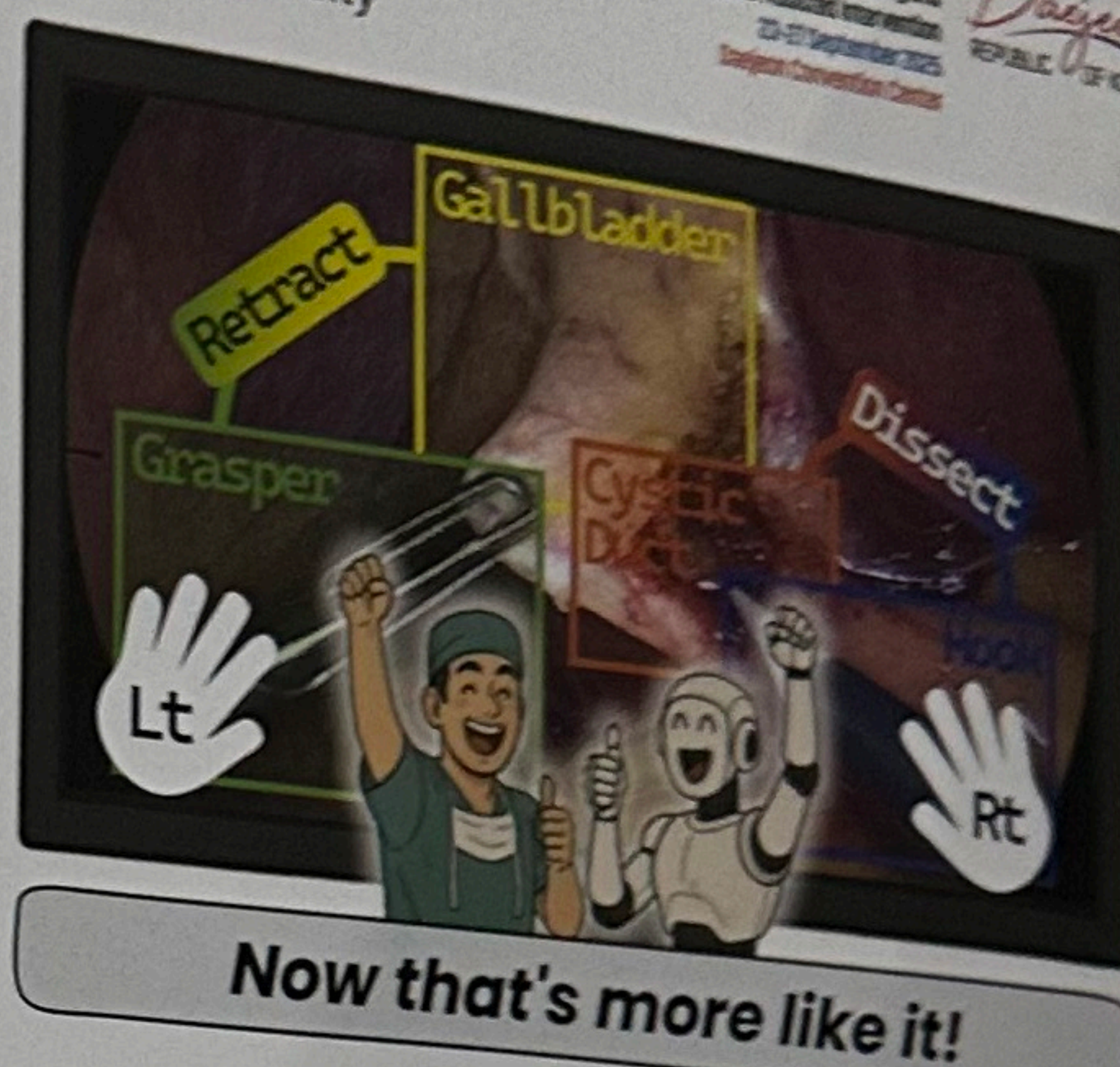
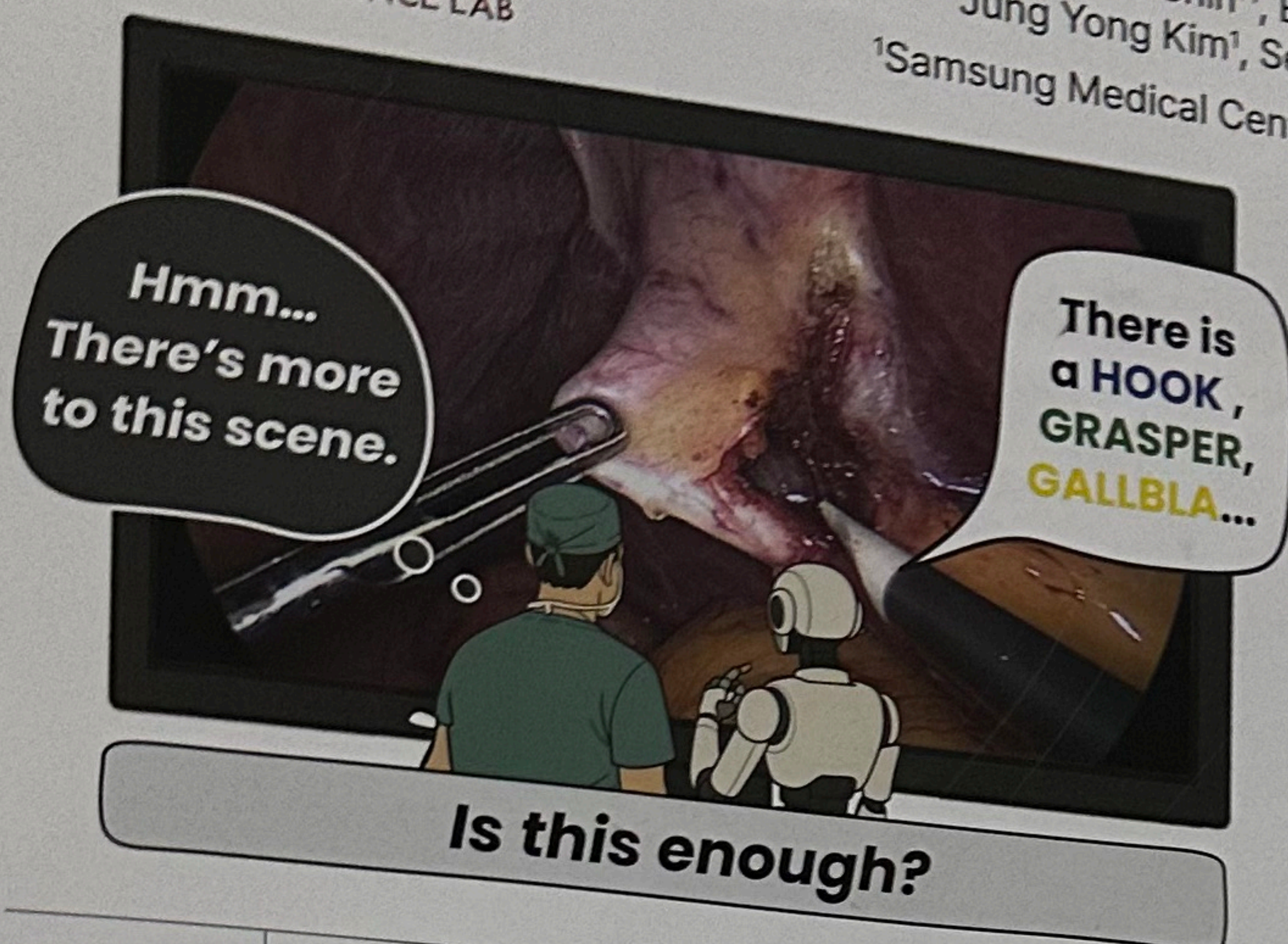
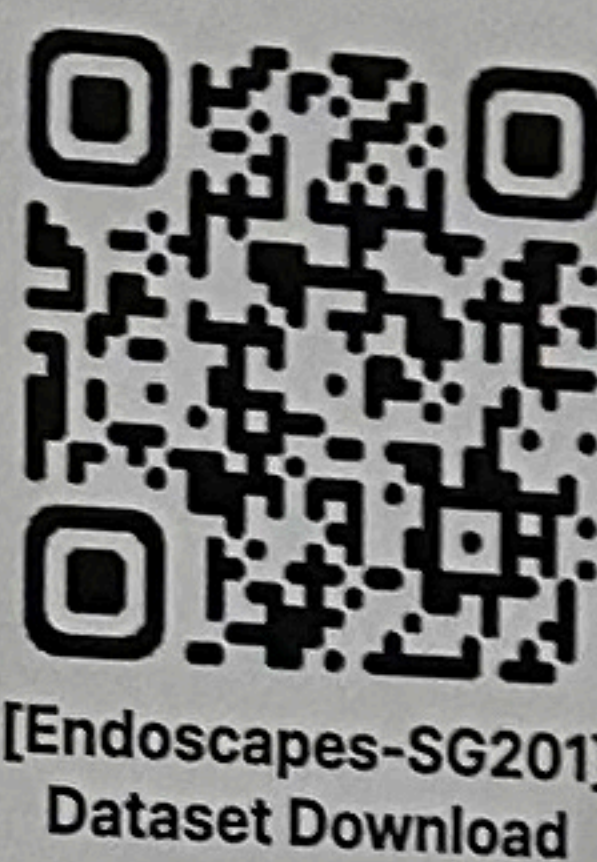


Towards Holistic Surgical Scene Graph

Jongmin Shin¹, Enki Cho^{1,2}, Ka Young Kim^{1,2},
 Jung Yong Kim¹, Seong Tae Kim², Namkee Oh¹
¹Samsung Medical Center (SMC), ²Kyung Hee University



Methods	Training Dataset	Contained Information				
		Surgical Tool	Spatial Relation	Anatomy	Action	Hand ID
Islam et al.	EndoVis-18	✓(9)	-	✓(1)	✓(12)	-
Holm et al.	CaDIS	✓(12)	-	✓(4)	-	✓(1)
LG-CVS	Endoscapes-Bbox201	✓(1)	✓(3)	✓(5)	-	-
SSG-Com (Ours)	Endoscapes-SG201 (Ours)	✓(6)	✓(3)	✓(5)	✓(6)	✓(3)



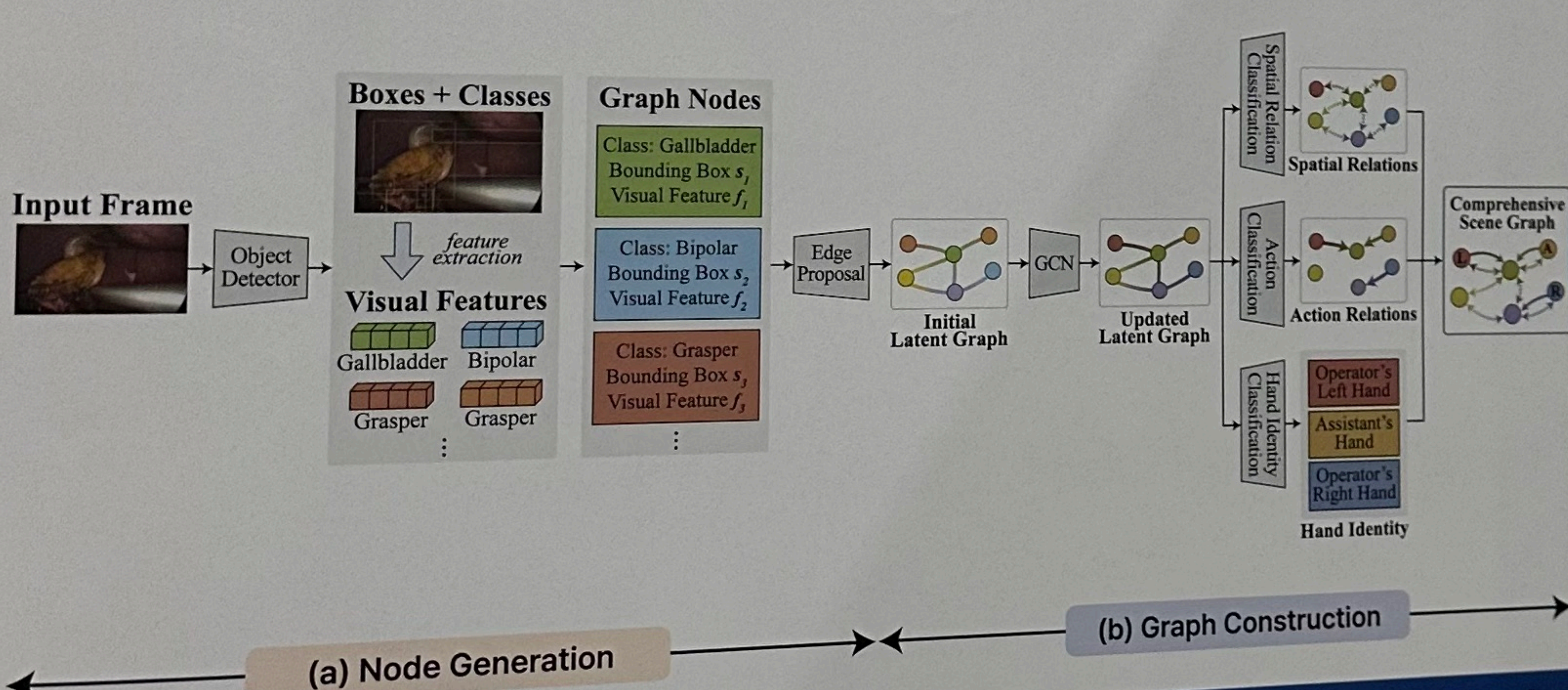
Endoscapes-SG201

A dataset providing unified annotations for bounding boxes, tool-action-target relationships, and hand identity.

SSG-Com

A graph-based model that integrates object locations, action relationships, and operator identity for a holistic scene understanding.

Methods



Integrated Learning of Action & Hand Identity

We enhance scene understanding by learning Action and Hand Identity jointly in an integrated manner within the scene graph.

(a) Node Generation

We generate the basic graph nodes by extracting RoI features from the bounding boxes of detected surgical tools and anatomical structures.

(b) Graph Construction

We learn a rich, comprehensive scene representation by simultaneously optimizing for edge relationships (Action) and node attributes (Hand Identity) within a single framework.

Quantitative Results

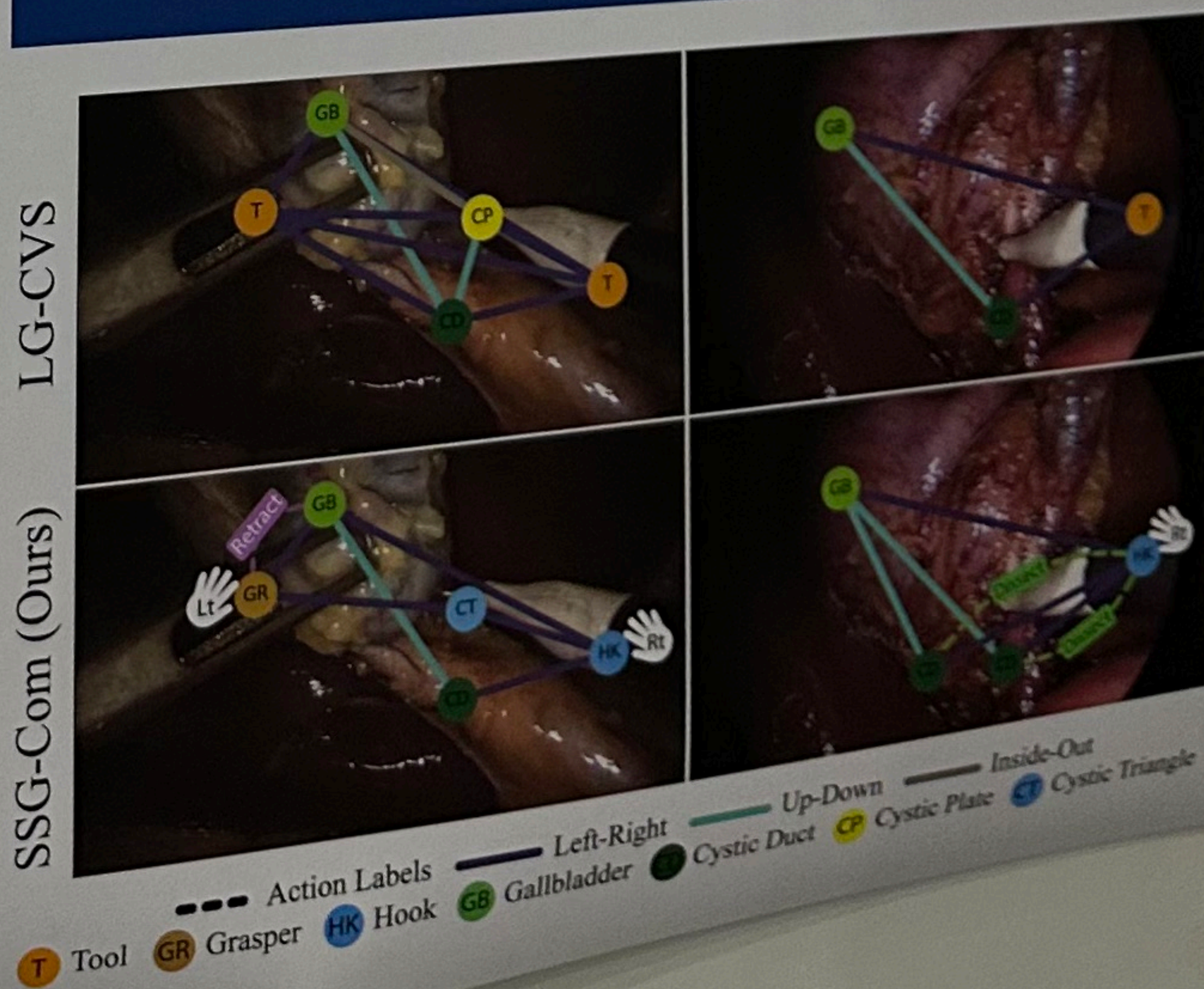
(a) Triplet Recognition Performance

Dataset	Method	Graph			Triplet (mAP)
		Spa.	Act.	Hand	
Endoscape-BBox201 [8]	ResNet50-DetInit				9.2
	LG-CVS	✓			13.8
Endoscape-SG201	ResNet50-DetInit				9.7
	LG-CVS	✓			18.0
	SSG-Com (Ours)	✓	✓		23.5
	SSG-Com (Ours)	✓	✓	✓	24.2

(b) CVS Performance

Dataset	Method	CVS (mAP)
Endoscape-BBox201 [8]	DeepCVS	42.4
	LayoutCVS	42.8
	ResNet50-DetInit	57.3
	LG-CVS	62.3
	DeepCVS	40.7
Endoscape-SG201	DeepCVS	44.0
	LayoutCVS	55.3
	ResNet50-DetInit	63.2
	SSG-Com (Ours)	64.6

Qualitative Results



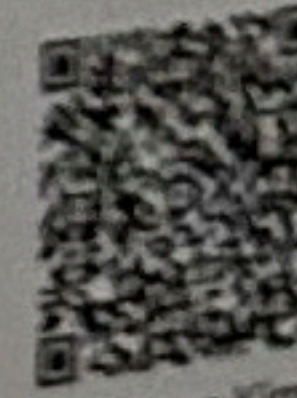
[About us]



Jongmin Shin



Enki Cho



Ka Young Kim