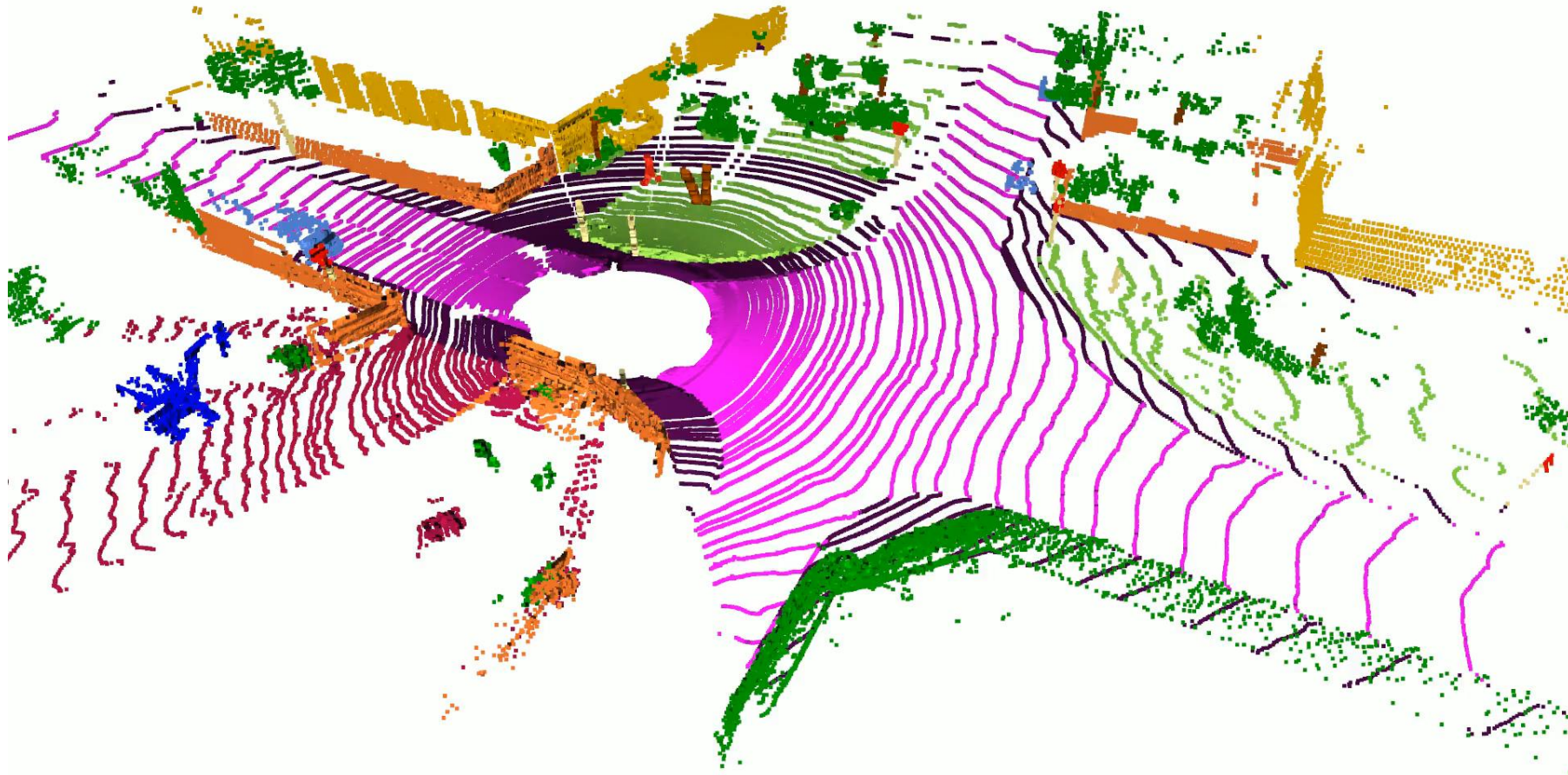


Temporal Consistent 3D LiDAR Representation Learning for Semantic Perception in Autonomous Driving

**Lucas Nunes, Louis Wiesmann, Rodrigo Marcuzzi, Xieyuanli Chen,
Jens Behley, Cyrill Stachniss**

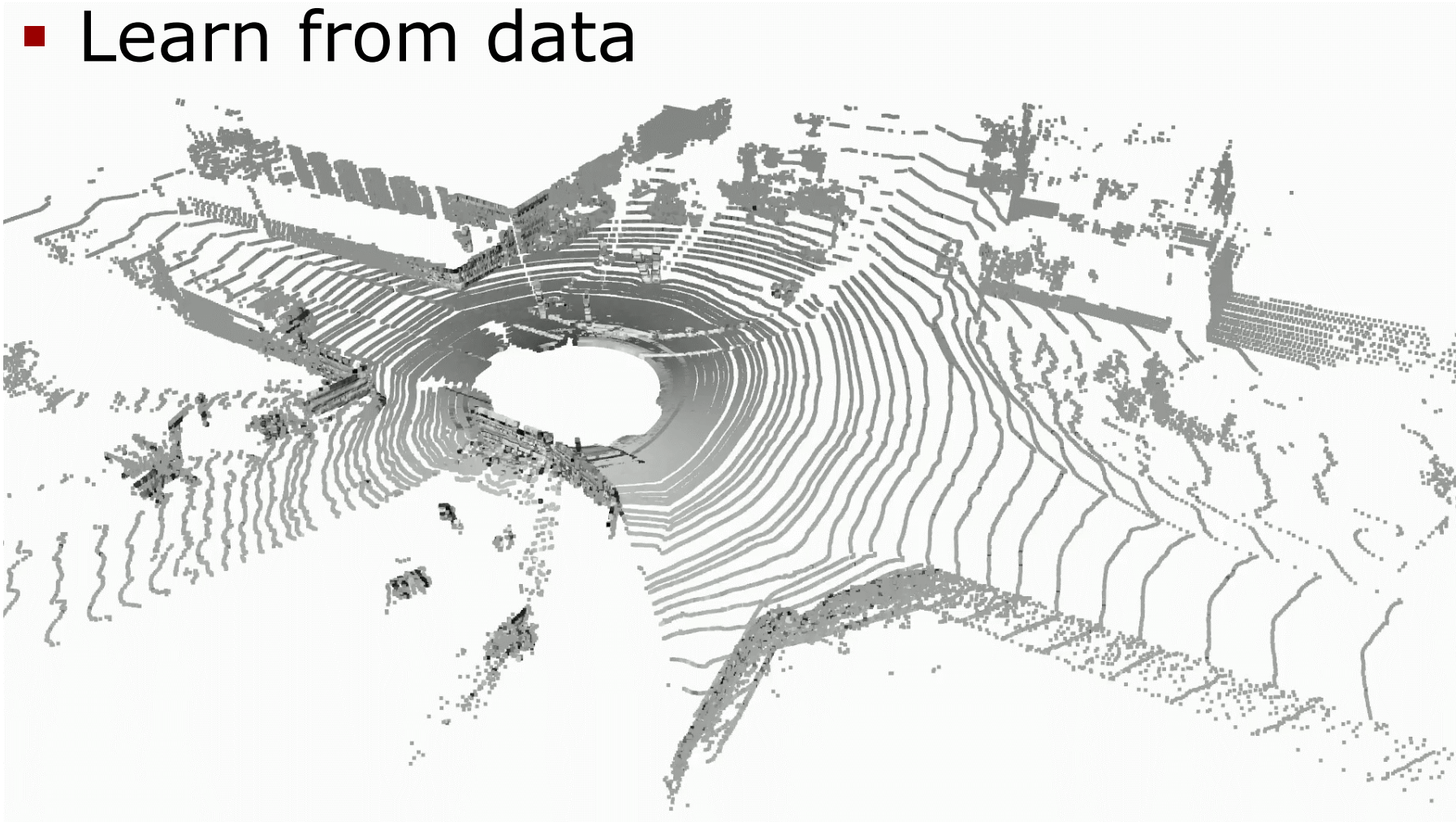
Motivation

- Scene understanding for self-driving
- Learning-based methods

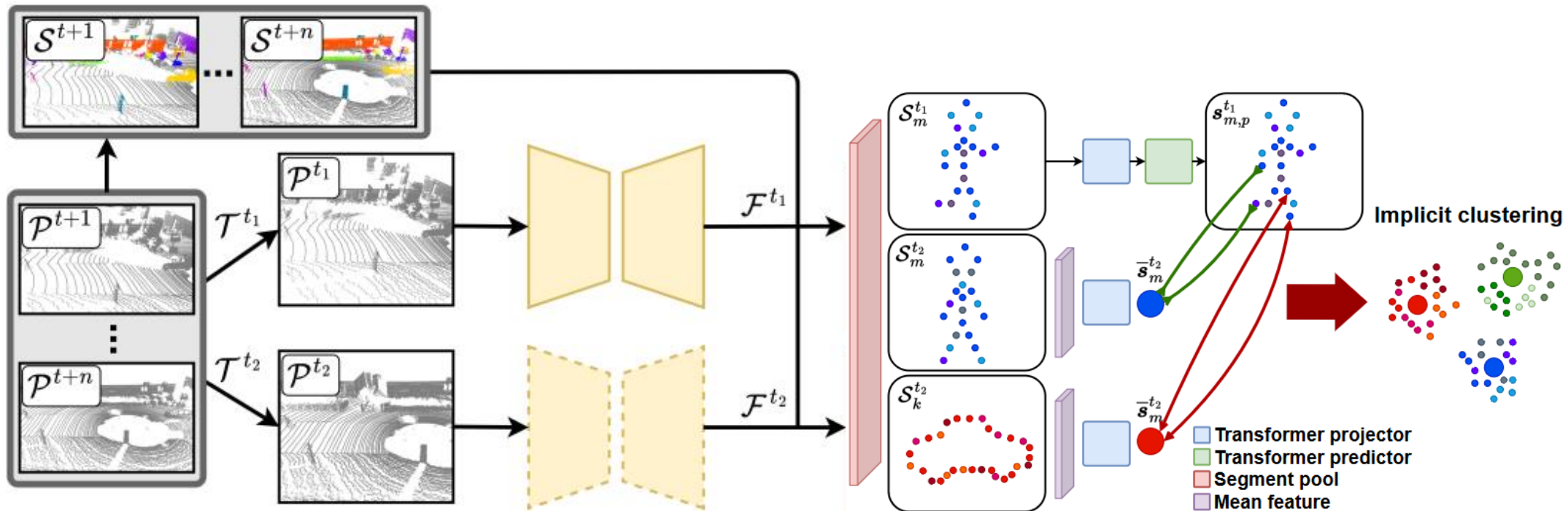


Motivation

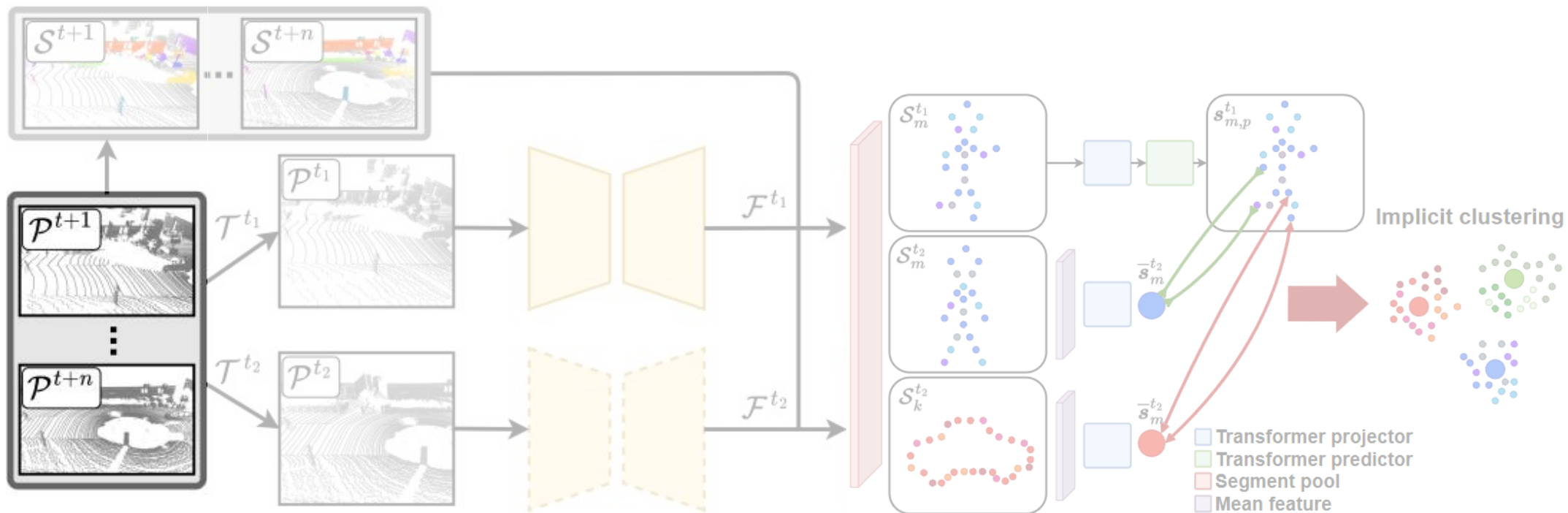
- Expensive data annotation
- Learn from data



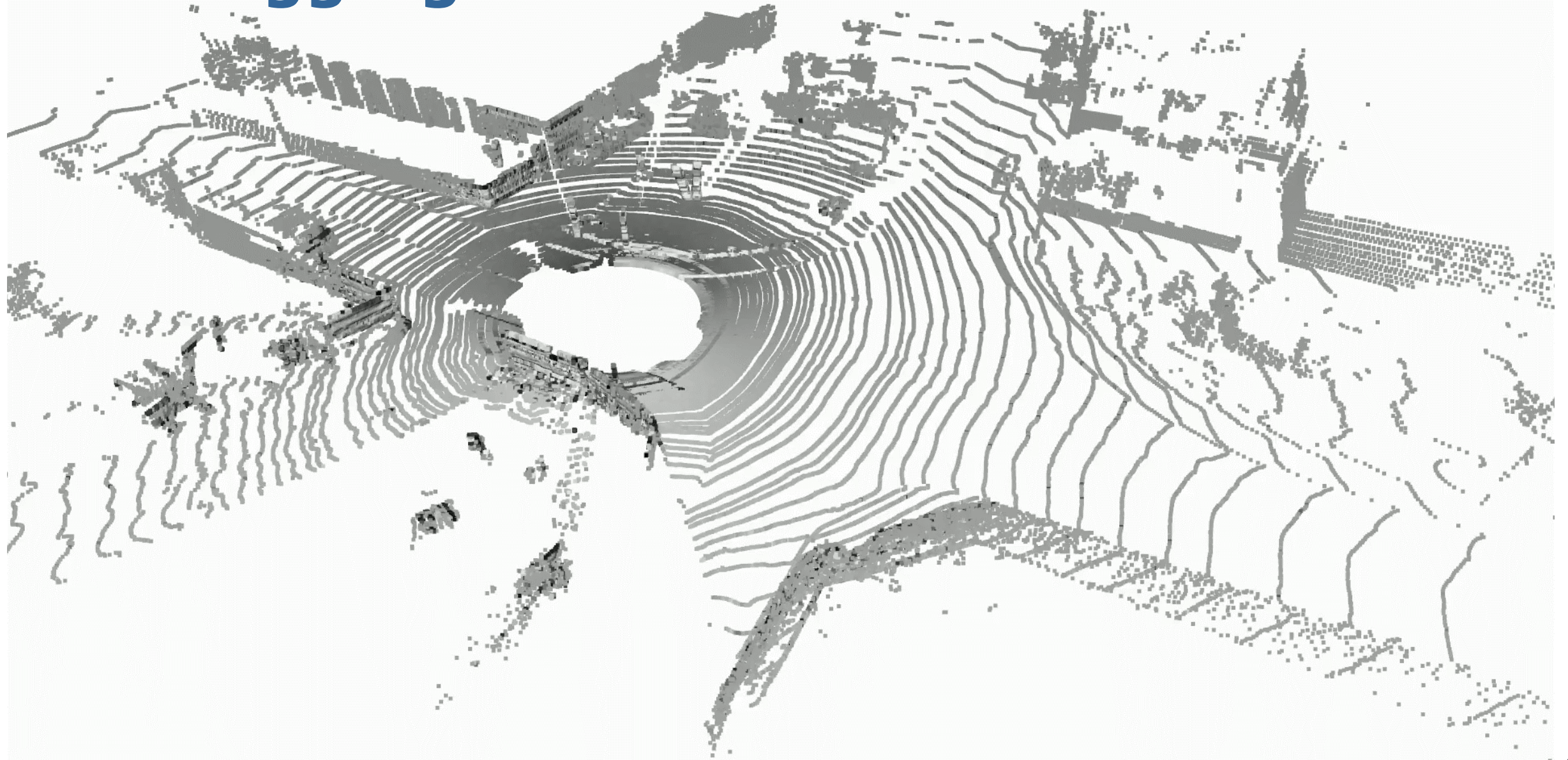
Our Approach (TARL)



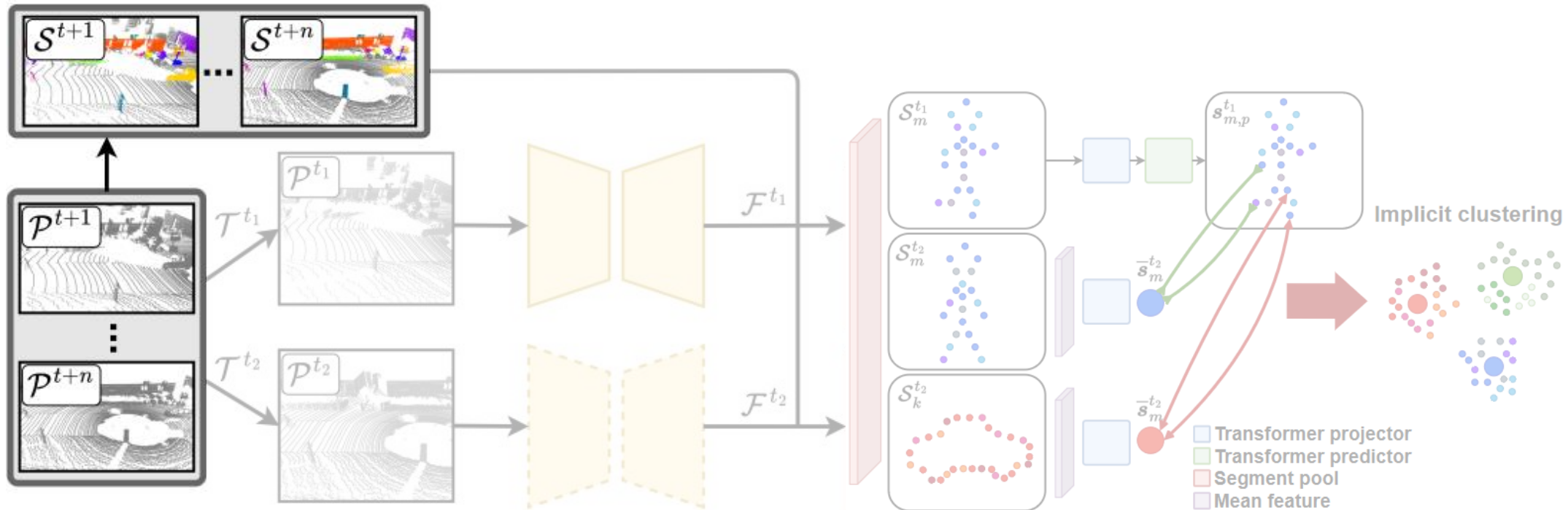
Scan Aggregation



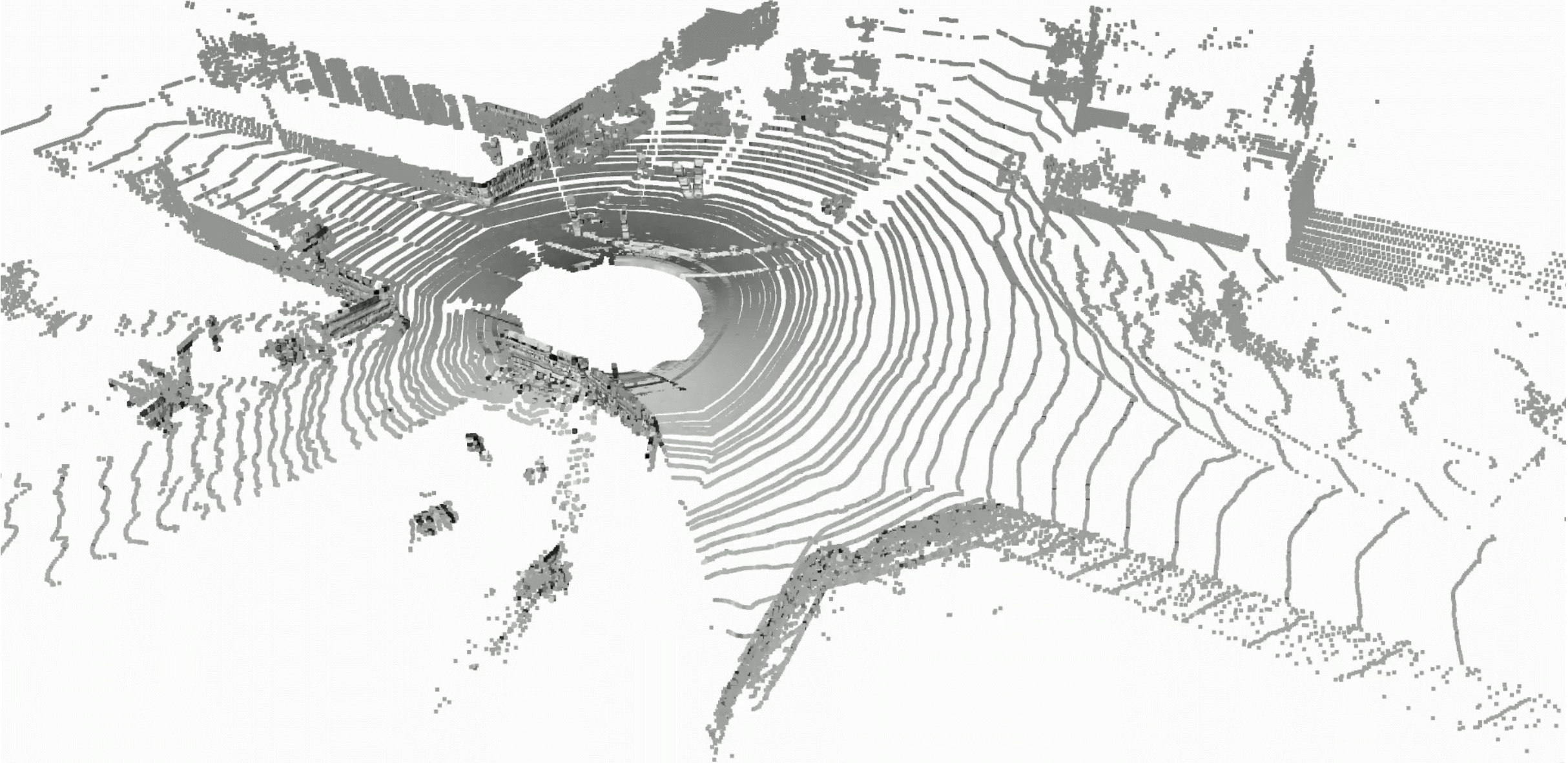
Scan Aggregation



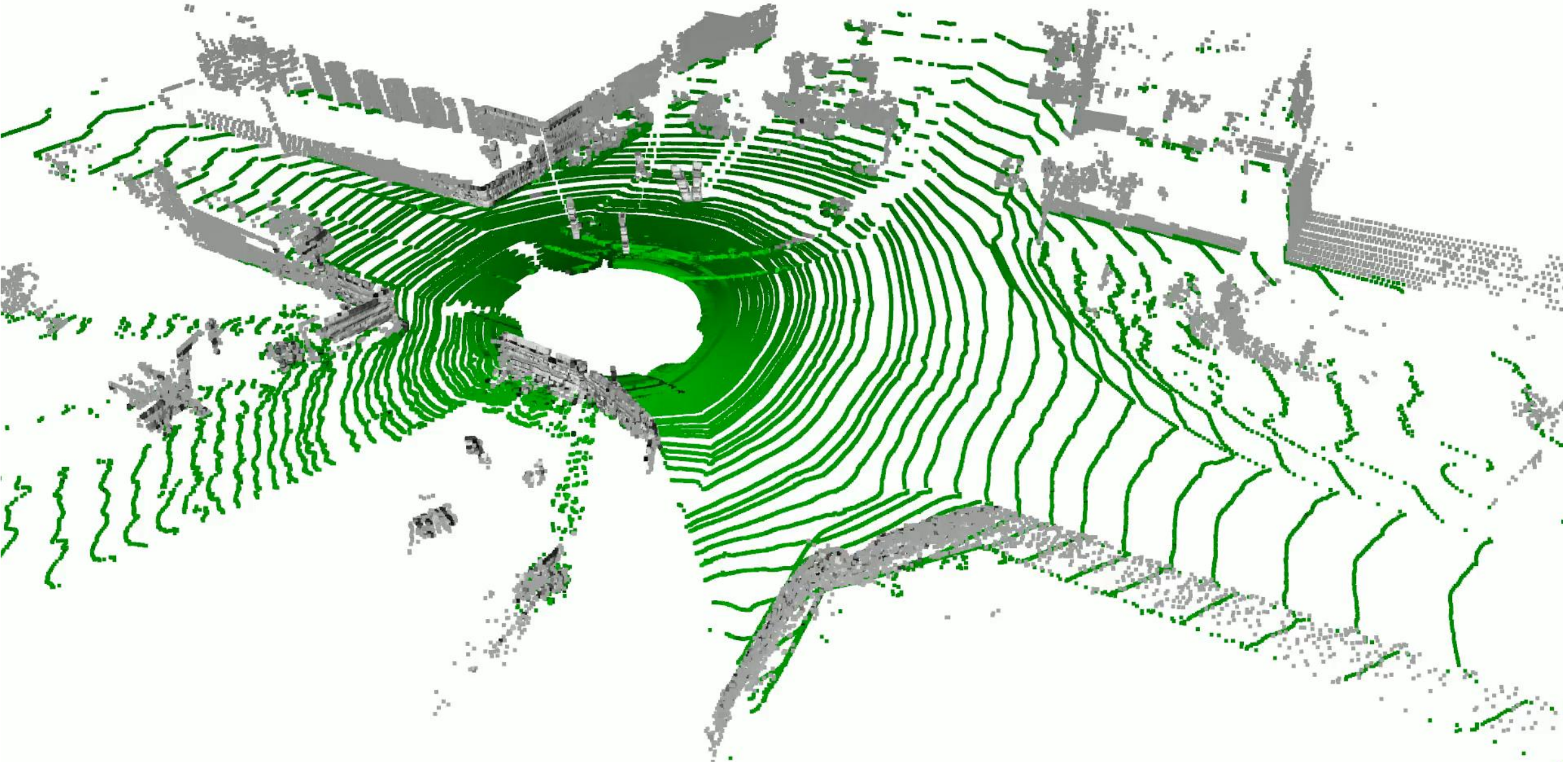
Temporal Views



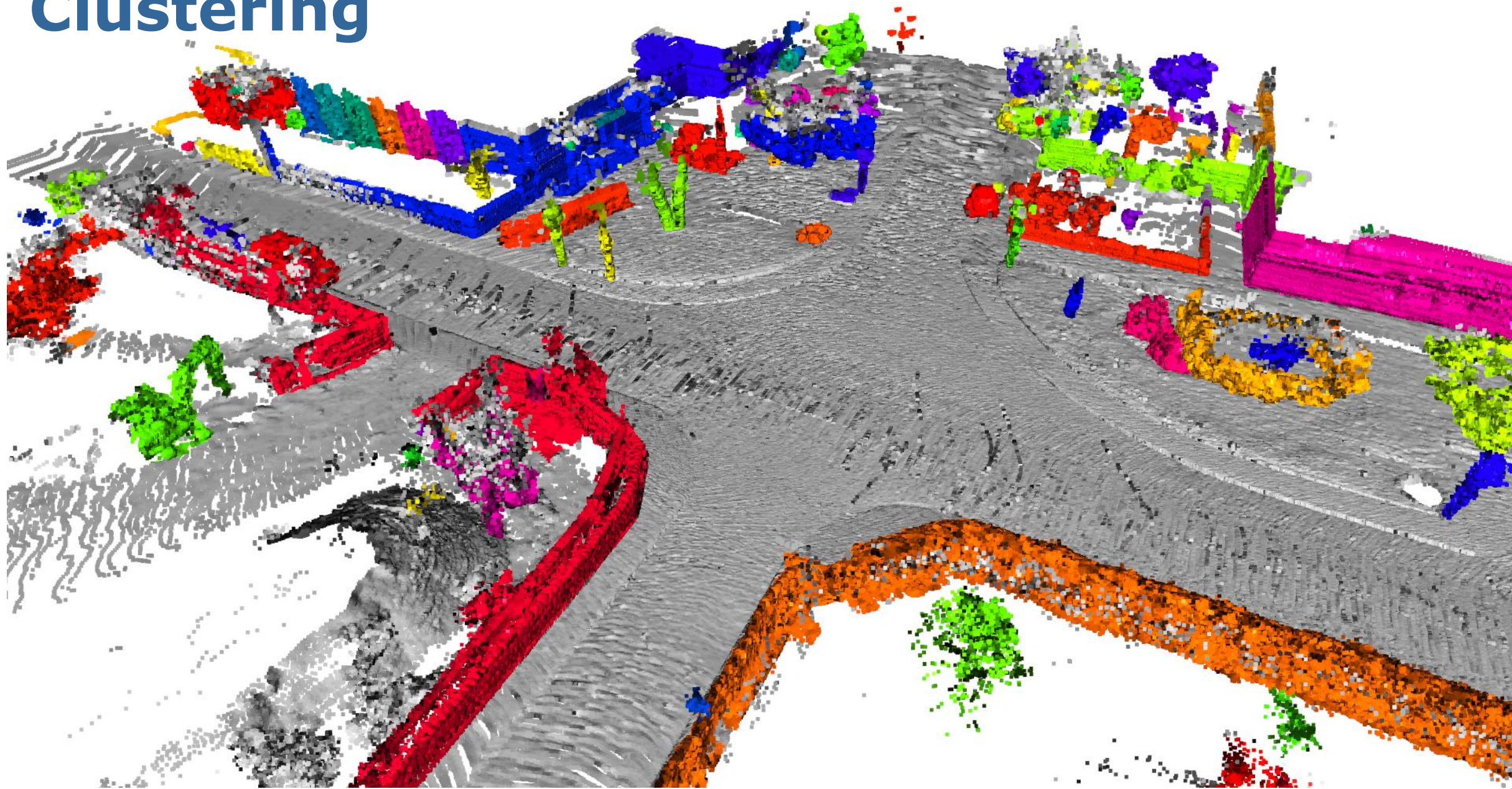
Ground Removal



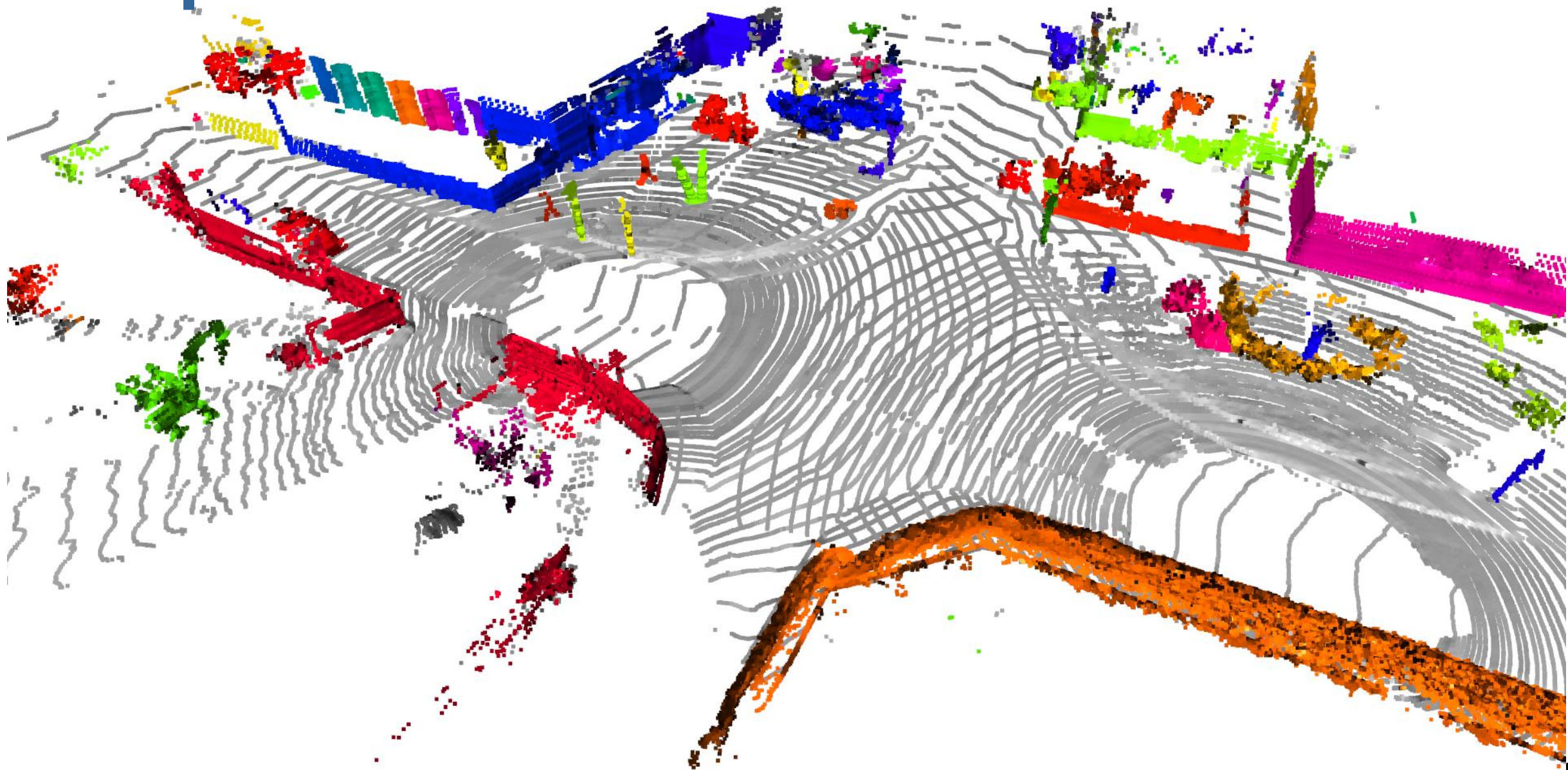
Ground Removal



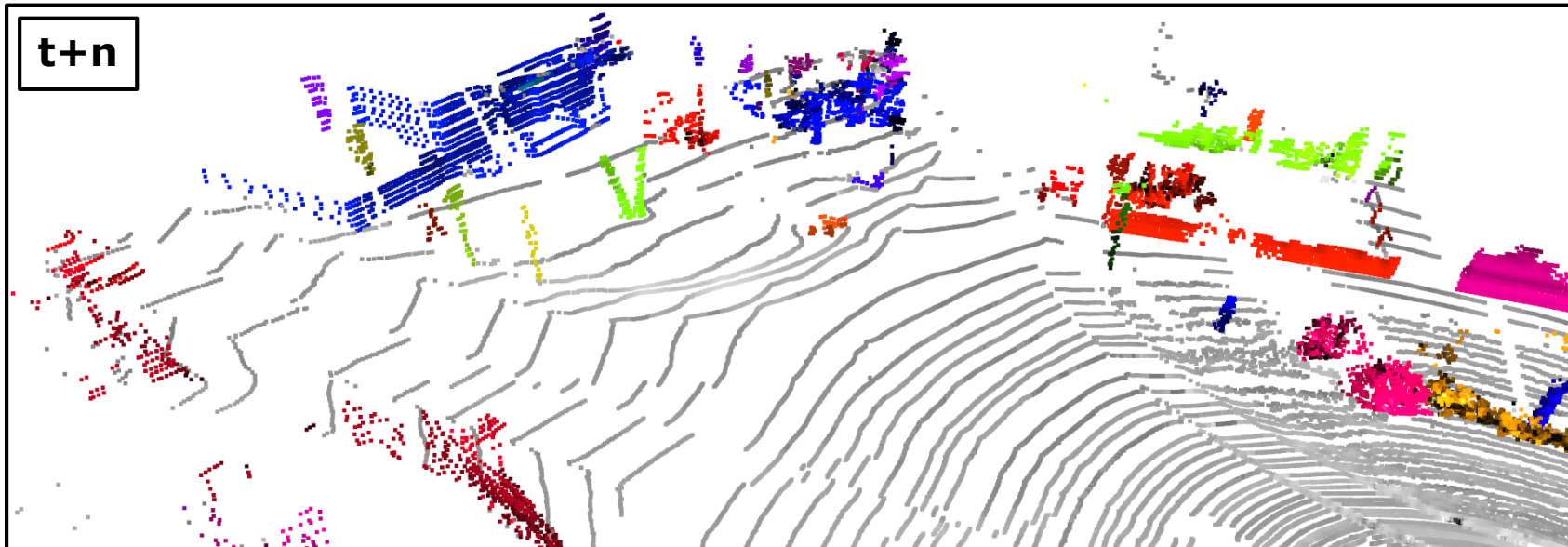
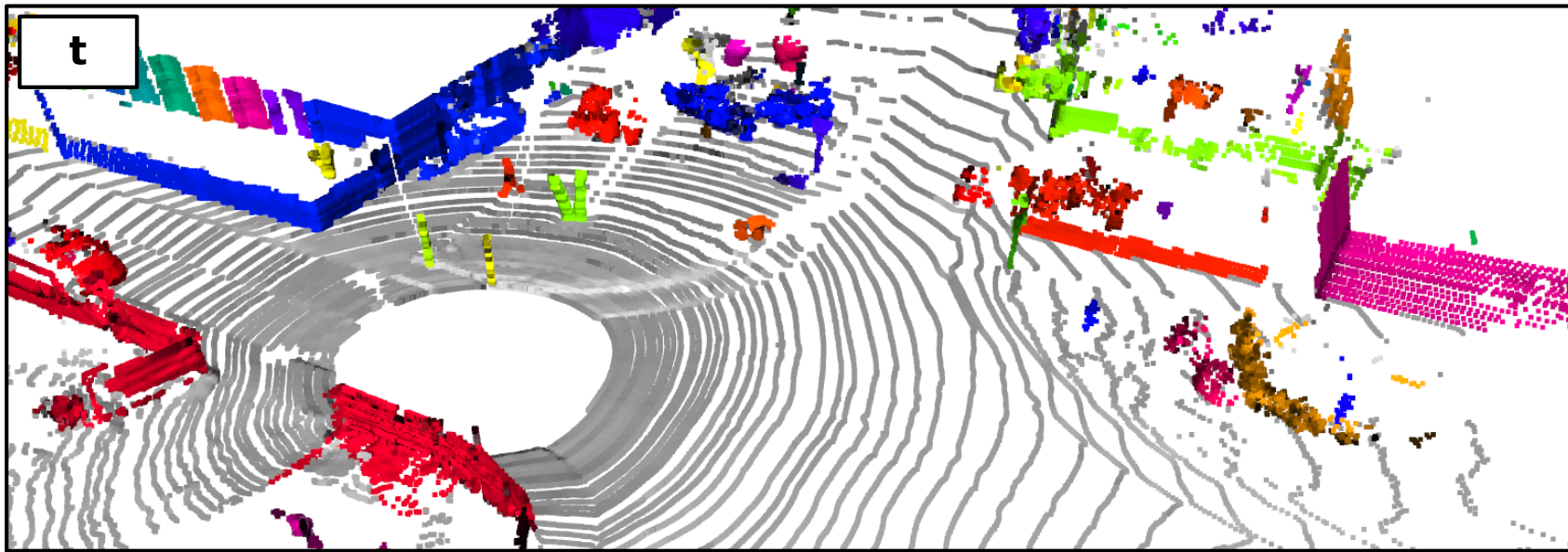
Clustering



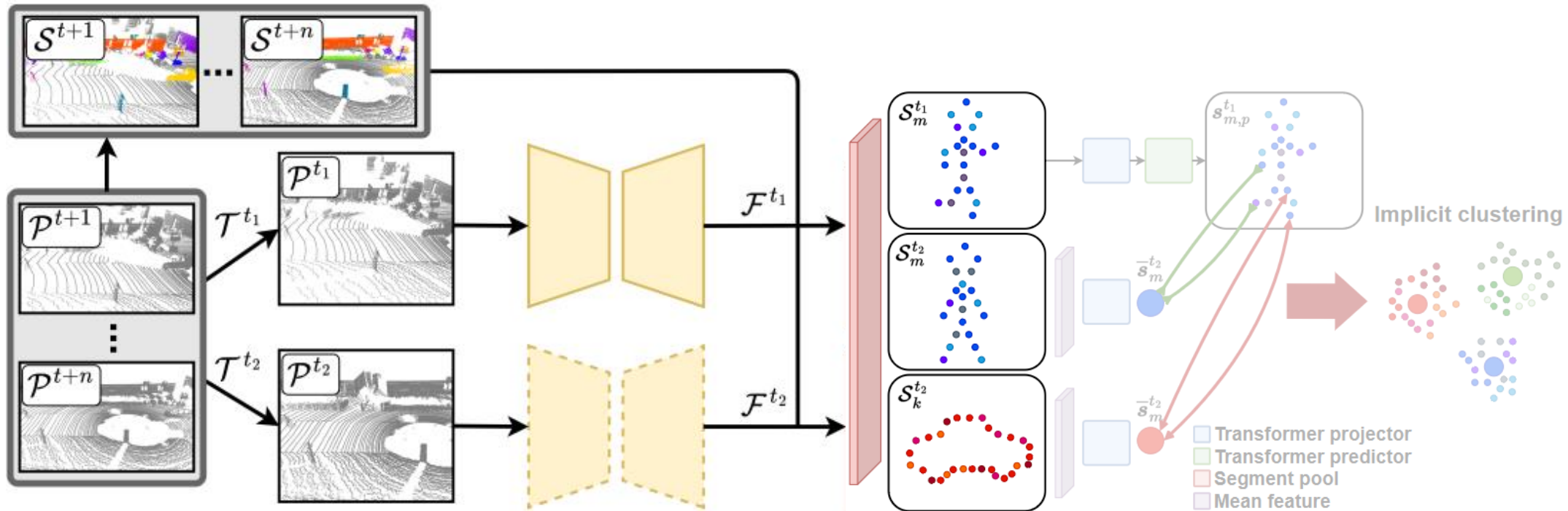
Temporal Views



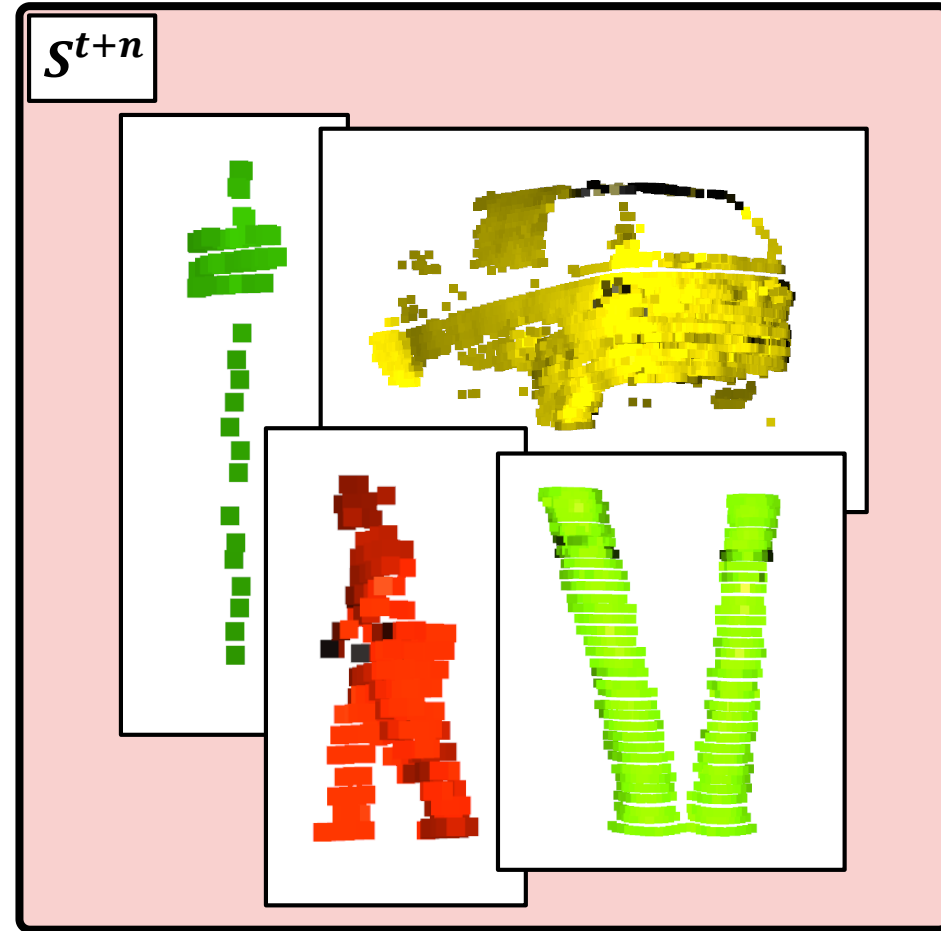
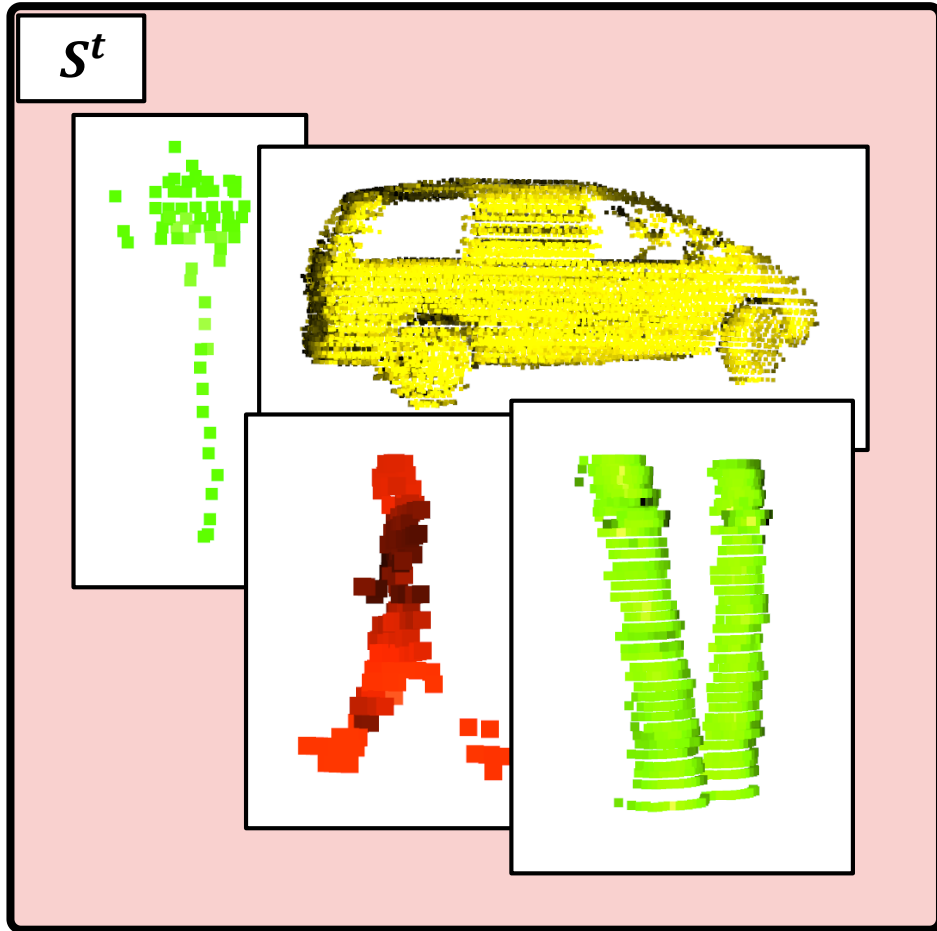
Temporal Views



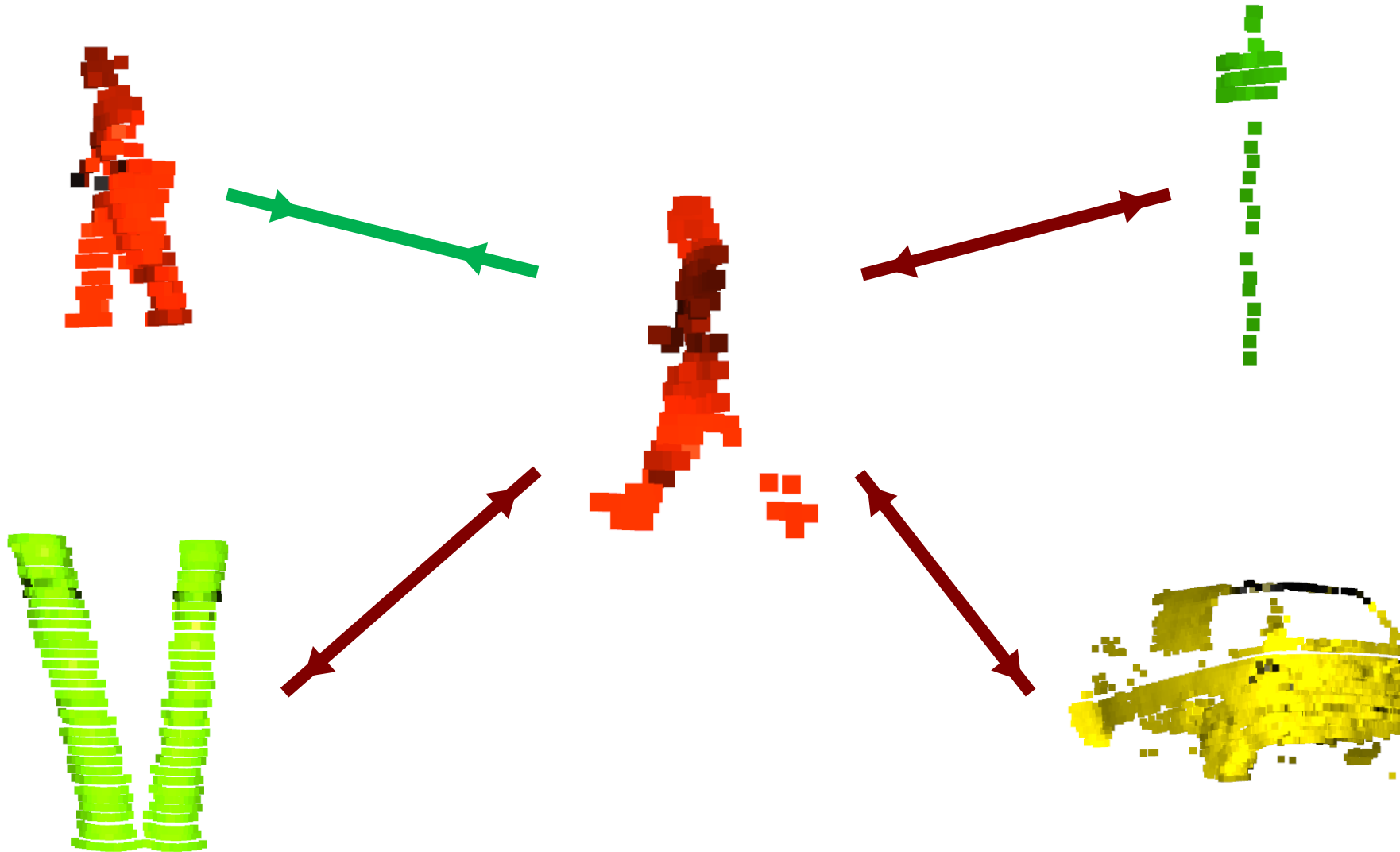
Temporal Segments Discrimination



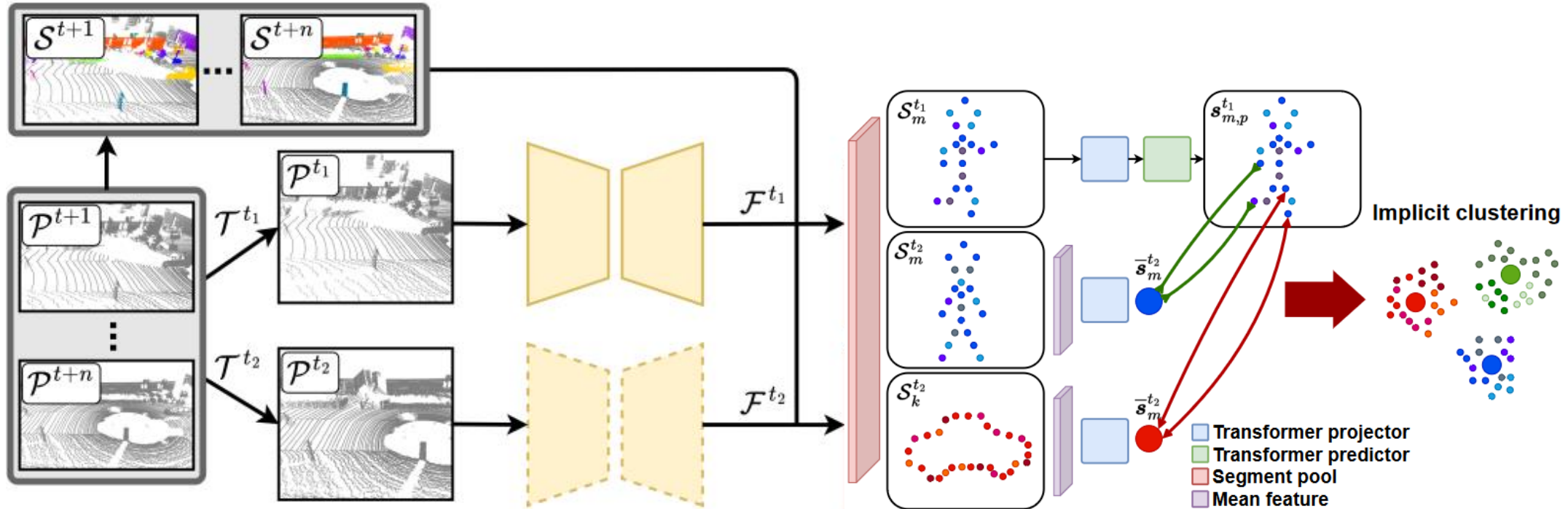
Segment Pooling



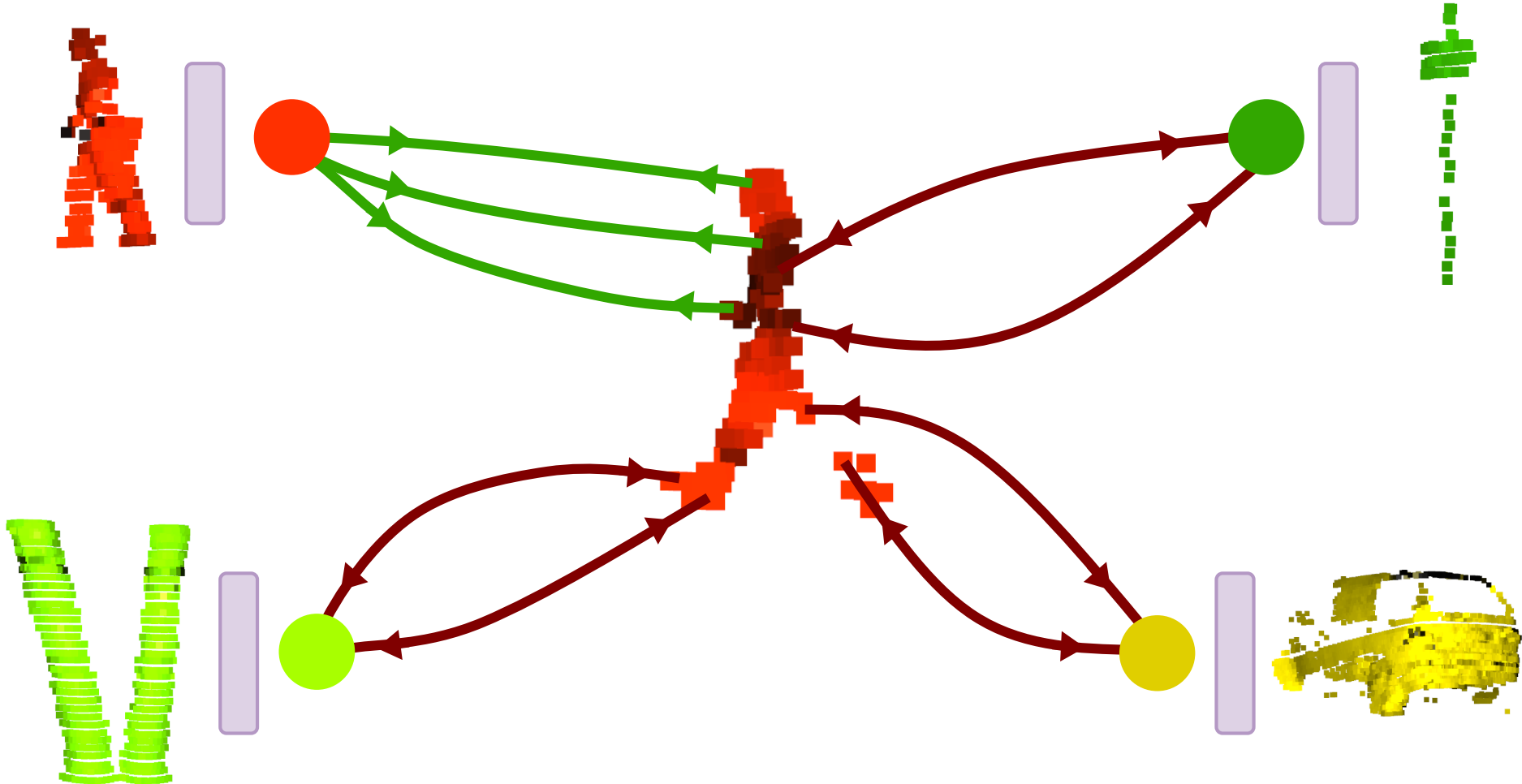
Temporal Segments Discrimination



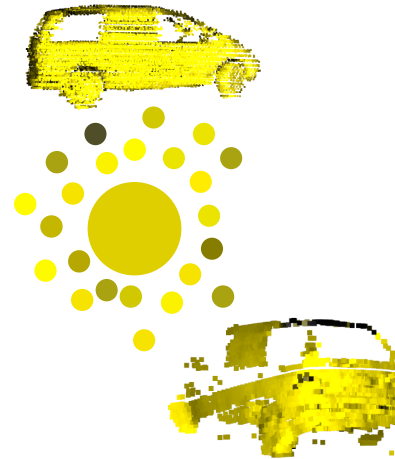
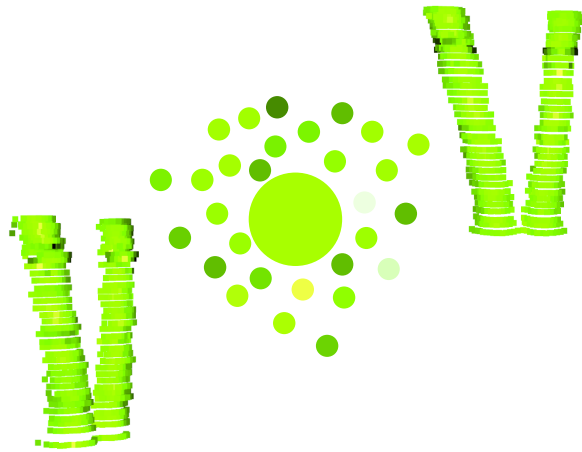
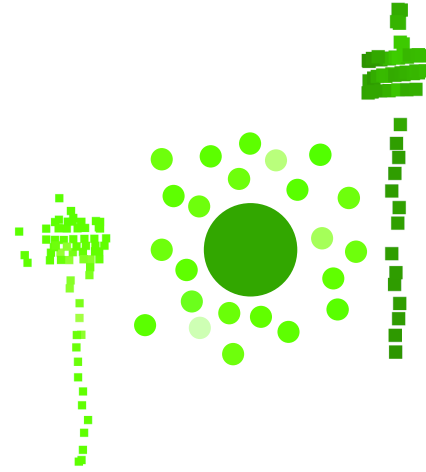
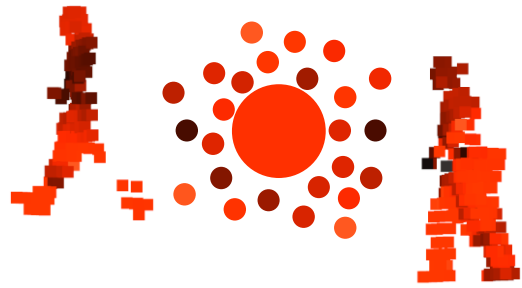
Implicit Clustering



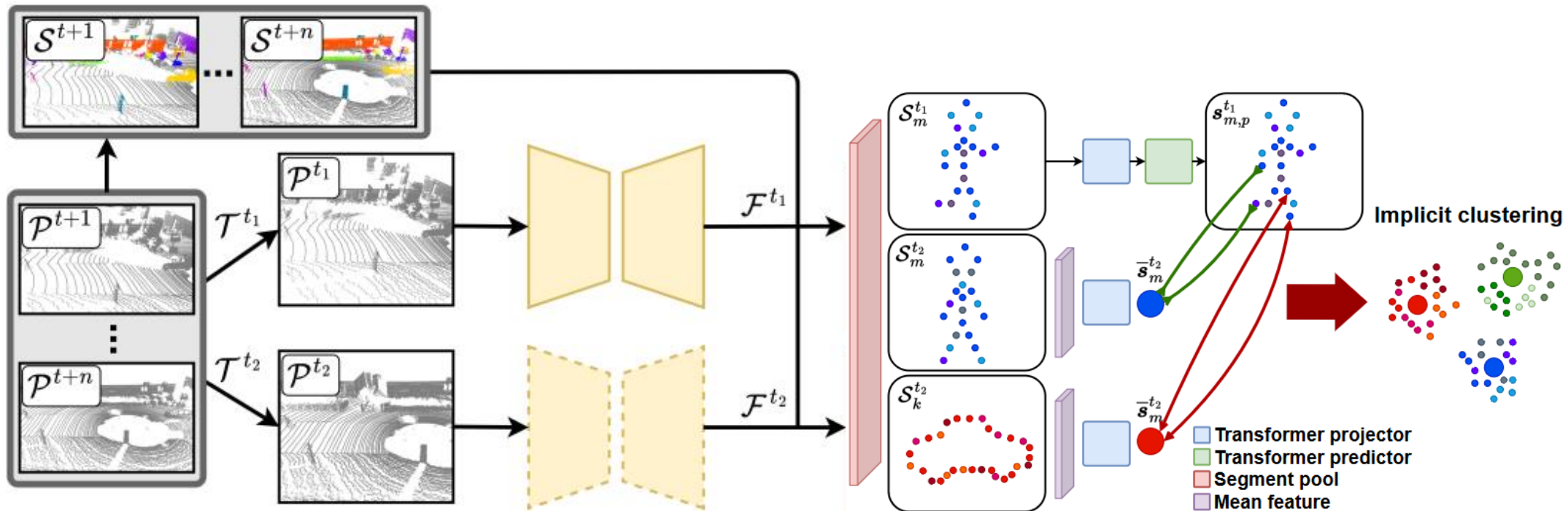
Implicit Clustering



Implicit Clustering



Our Approach (TARL)



Less Labels and Better Results

Pre-training and fine-tuning on SemantickITTI for semantic segmentation

Method	Scribbles	0.1%	1%	10%	50%	100%
Scratch	54.96	29.35	42.77	53.96	58.27	59.03
PointContrast	54.52	32.63	44.62	58.68	59.98	61.45
DepthContrast	55.90	31.66	48.05	57.11	60.99	61.14
SegContrast	56.70	32.75	44.83	56.31	60.45	61.02
TARL (Ours)	57.25	38.59	51.42	60.34	61.42	61.47

10% labels

Less Labels and Better Results

Pre-training and fine-tuning on SemanticKITTI for panoptic segmentation

	10%		50%		100%	
Method	PQ	IoU	PQ	IoU	PQ	IoU
Scratch	47.20	53.53	55.32	61.94	55.40	59.75
PointContrast	47.57	54.63	54.21	59.48	55.85	61.49
DepthContrast	46.85	51.27	54.55	59.60	56.15	60.81
SegContrast	47.02	53.47	55.38	60.04	56.73	61.96
TARL (Ours)	51.27	57.59	56.10	62.36	56.57	62.05

Better Than Supervised Pre-Training

Pre-training on SemanticKITTI and fine-tuning to nuScenes for semantic segmentation

Method	Mini	Full
Scratch	26.94	66.03
Supervised pre-training	38.39	67.35
PointContrast	31.92	67.31
DepthContrast	27.81	64.70
SegContrast	31.27	67.70
TARL (Ours)	39.36	68.26

Better Than Supervised Pre-Training

Pre-training on SemanticKITTI and fine-tuning to nuScenes for panoptic segmentation

	Mini		Full	
	PQ	IoU	PQ	IoU
Scratch	23.78	23.96	52.98	58.17
Supervised pre-training	24.77	23.60	53.19	58.05
PointContrast	26.58	25.46	51.06	56.39
DepthContrast	28.66	27.30	51.51	57.06
SegContrast	28.84	26.79	52.31	57.24
TARL (Ours)	32.22	30.73	53.26	59.14



Summary

- Representation learning method for LiDAR data
- Learn a representation consistent across time
- Requires 1/10 train data for semantic segmentation
- Better performance than supervised pre-training



Code



Paper



Code



Paper