



# TRANSFER YOUR PERSPECTIVE:

## Controllable 3D Generation from Any Viewpoint in a Driving Scene



Tai-Yu Pan<sup>1</sup>, Sooyoung Jeon<sup>1</sup>, Mengdi Fan<sup>1</sup>, Jinsu Yoo<sup>1</sup>, Zhenyang Feng<sup>1</sup>, Mark Campbell<sup>2</sup>, Kilian Q. Weinberger<sup>2</sup>, Bharath Hariharan<sup>2</sup>, Wei-Lun Chao<sup>1</sup>

<sup>1</sup>The Ohio State University

<sup>2</sup>Cornell University

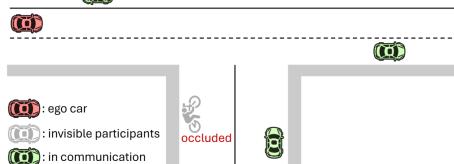
TYP (author)

TYP (project)

### Highlights

- Introduce a **new research direction**: generating realistic sensor data from arbitrary viewpoints to enable collaborative perception development
- Propose TYP**: the first solution, successfully transferring single-agent Waymo to collaborative ColWaymo
- Unlock large-scale pre-training** for collaborative perception with **semi-real data**

### Future Autonomous Driving



**Collaborative Driving:**  
safer AD system by communicating  
perception signals or results

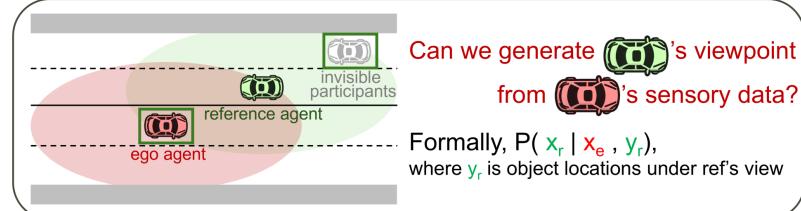
However, Challenges of  
Real-World Data Collection for  
Collaborative Perception Development:

- Nx more effort than ego-centric setups
- Hard to coordinate multiple vehicles on the same road, same time, within limited range

Thus, existing datasets:

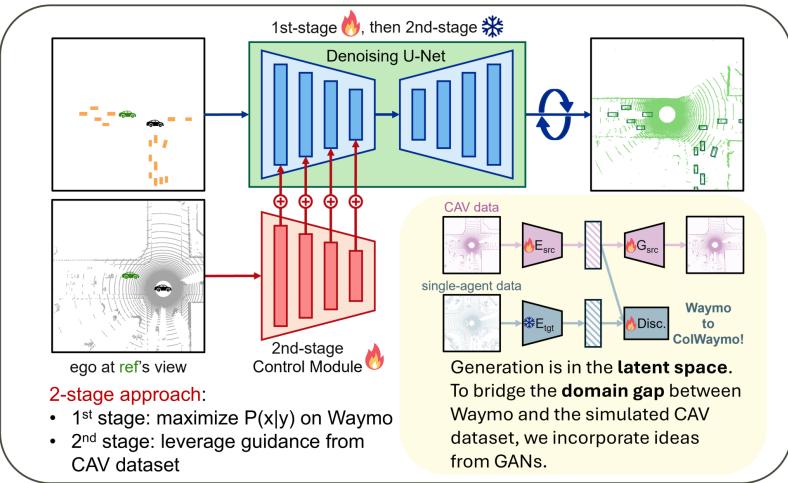
- Primarily simulated
- Few agents (typically only 2)

### New Research Direction

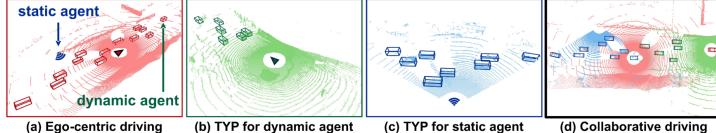


**Problem:** We want to convert Waymo into a collaborative dataset—but it only contains ego-view data ( $x_e$ ), and lacks reference-view data ( $x_r$ )!

### TRANSFER YOUR PERSPECTIVE (TYP)



### Qualitative Results

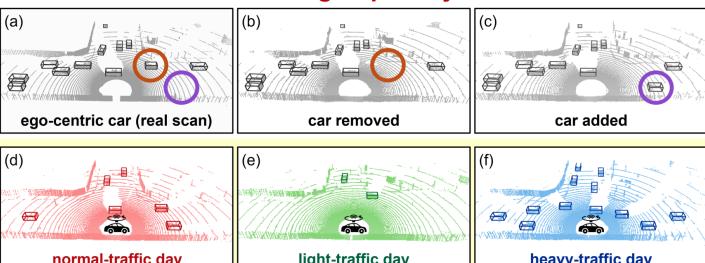


### Unlock CAV Pre-Training w/ ColWaymo

Pre-training: scratch / OPV2V / ColWaymo, fine-tuning: V2V4Real, test: V2V4Real

Method	Pre-Train	FT. on 0 Scene			FT. on 5 Scene			FT. on 10 Scene			FT. on 32 Scene						
		s	m	l	s	m	l	s	m	l	s	m	l				
No Fusion	ego's gt only	—	—	—	0.44	0.16	0.06	0.29	0.66	0.29	0.08	0.42	0.71	0.34	0.08	0.47	
Early Fusion [9]	OPV2V	<b>0.54</b>	0.20	0.07	0.31	0.32	0.18	0.16	0.30	0.73	0.28	0.24	0.48	0.79	0.42	0.45	0.60
ColWaymo (ours)	OPV2V	<b>0.50</b>	<b>0.24</b>	<b>0.24</b>	<b>0.35</b>	<b>0.68</b>	<b>0.34</b>	<b>0.32</b>	<b>0.51</b>	<b>0.78</b>	<b>0.39</b>	<b>0.33</b>	<b>0.57</b>	<b>0.83</b>	<b>0.48</b>	<b>0.50</b>	<b>0.65</b>
ColWaymo (ours)	scratch	0.44	0.18	0.16	0.31	0.32	0.18	0.15	0.25	0.73	0.32	0.31	0.50	0.80	0.47	0.54	0.65
Late Fusion [80]	OPV2V	<b>0.60</b>	<b>0.27</b>	<b>0.28</b>	<b>0.44</b>	0.55	0.26	0.38	0.42	0.73	0.42	<b>0.51</b>	0.58	0.79	0.53	0.56	0.67
ColWaymo (ours)	OPV2V	0.40	0.19	0.15	0.25	<b>0.60</b>	<b>0.28</b>	<b>0.44</b>	<b>0.47</b>	<b>0.76</b>	<b>0.43</b>	<b>0.51</b>	<b>0.61</b>	<b>0.82</b>	<b>0.58</b>	<b>0.61</b>	<b>0.71</b>
AttFuse [80]	OPV2V	0.51	0.19	0.05	0.31	0.54	0.22	0.11	0.37	0.77	0.40	0.21	0.54	0.83	0.53	0.40	0.65
ColWaymo (ours)	OPV2V	<b>0.66</b>	<b>0.35</b>	<b>0.11</b>	<b>0.45</b>	<b>0.65</b>	<b>0.29</b>	<b>0.16</b>	<b>0.46</b>	<b>0.83</b>	<b>0.48</b>	<b>0.33</b>	<b>0.61</b>	<b>0.88</b>	<b>0.58</b>	<b>0.53</b>	<b>0.72</b>
ColWaymo (ours)	scratch	0.40	0.15	0.12	0.31	0.70	0.28	0.17	0.43	0.81	0.49	0.30	0.61	0.88	0.55	0.49	0.61
V2X-ViT [79]	OPV2V	0.51	0.24	0.07	0.33	0.48	0.23	0.16	0.35	0.76	0.38	0.22	0.53	0.81	0.49	0.35	0.61
ColWaymo (ours)	OPV2V	<b>0.60</b>	<b>0.28</b>	<b>0.10</b>	<b>0.34</b>	<b>0.66</b>	<b>0.28</b>	<b>0.22</b>	<b>0.46</b>	<b>0.79</b>	<b>0.48</b>	<b>0.26</b>	<b>0.58</b>	<b>0.84</b>	<b>0.57</b>	<b>0.44</b>	<b>0.67</b>

### Editing Capability



Acknowledgement: NSF (IIS-2107077, IIS-2107161), Ohio Supercomputer Center