# Uncertainty-Aware Vehicle Orientation Estimation for Joint Detection-Prediction Models

Henggang Cui, Fang-Chieh Chou, Jake Charland,
Carlos Vallespi-Gonzalez, Nemanja Djuric

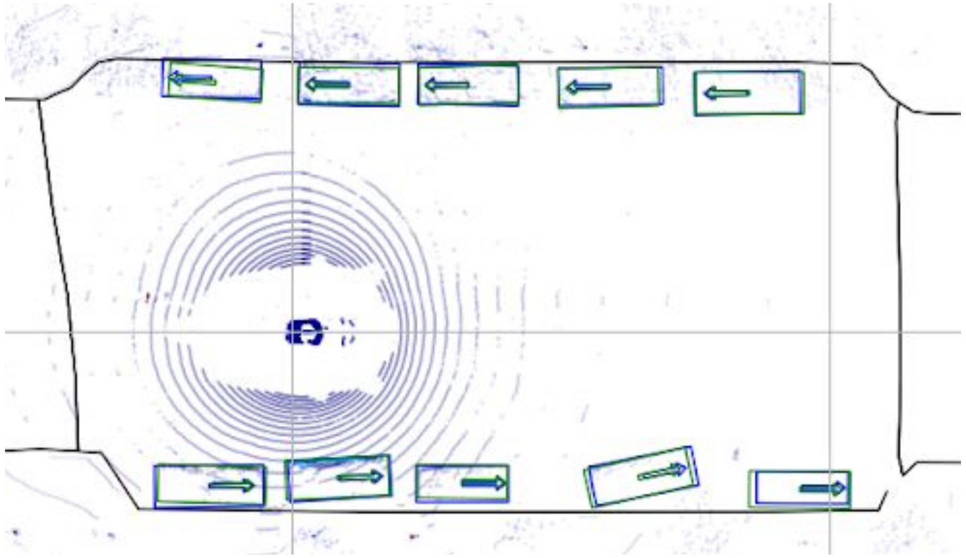
Uber Advanced Technologies Group

UBER

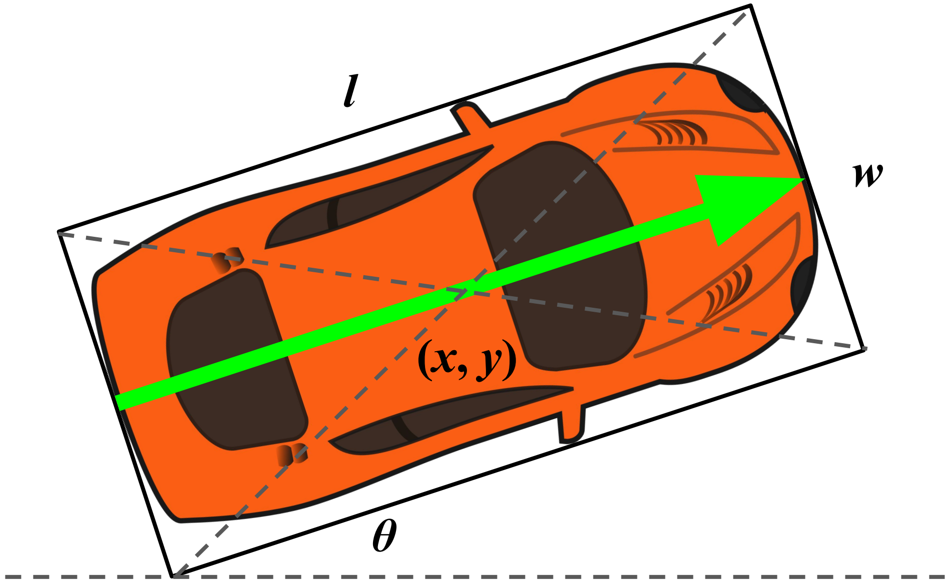
ATG

# Vehicle orientation estimation for autonomous driving

- Important property for downstream modules of an autonomous system, particularly for motion prediction of stationary or reversing actors where current methods struggle
- Assume the model produces detection bounding boxes from LiDAR points, as $(x, y, l, w, \theta)$
- Our work proposes a novel uncertainty-aware method to learn $\theta$



LiDAR points are the input, and detections output of our model

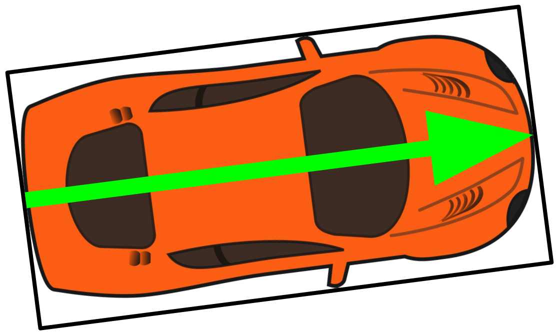


# Traditional orientation estimation method

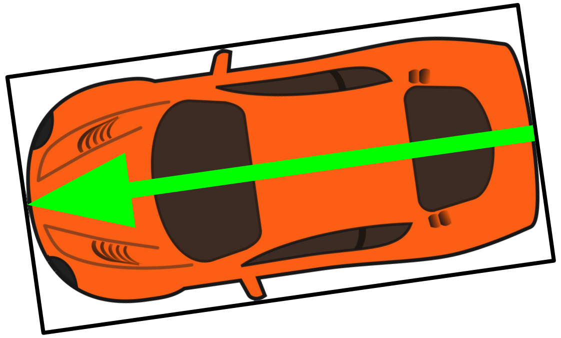
- Have the model output $\sin(\theta), \cos(\theta)$ and train them with

$$\mathcal{L}_{full} = \ell_1\big(\sin(\hat{\theta}) - \sin(\theta)\big) + \ell_1\big(\cos(\hat{\theta}) - \cos(\theta)\big)$$

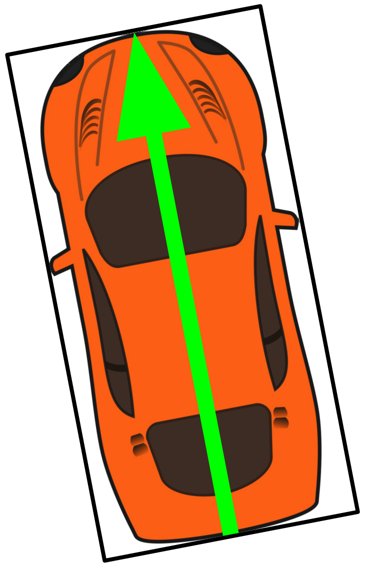
- However, a 180°-flipped orientation gets a higher loss than a 90°-error of the orientation
- This is empirically shown to hurt overall detection performance



**Ground-truth**



**180°-flipped orientation**
OK detection but high loss



**90°-error orientation**
Terrible detection but lower loss

# Flip-aware orientation estimation

- In addition to $\sin(\theta)$, $\cos(\theta)$, the model also outputs a flipped probability $p_f$, trained with

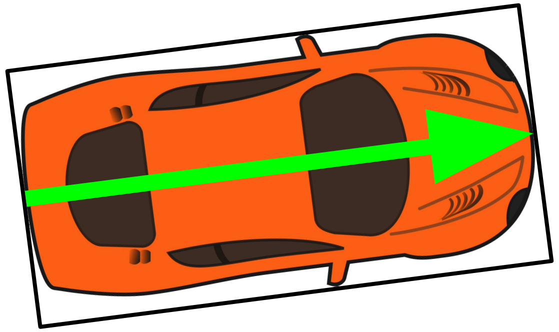
$$\mathcal{L}_{final} = \mathcal{L}_{half} + \min(\mathcal{L}_{full}, \mathcal{L}_{flipped}) + \text{CrossEntropy}(\hat{p}_f, \mathbb{1}_{\mathcal{L}_{full} > \mathcal{L}_{flipped}})$$

where $\mathcal{L}_{half} = \ell_1\big(\sin(2\hat{\theta}) - \sin(2\theta)\big) + \ell_1\big(\cos(2\hat{\theta}) - \cos(2\theta)\big)$ ; $\sin(2\hat{\theta}) = 2\sin(\hat{\theta})\cos(\hat{\theta})$ ; $\cos(2\hat{\theta}) = \cos^2(\hat{\theta}) - \sin^2(\hat{\theta})$

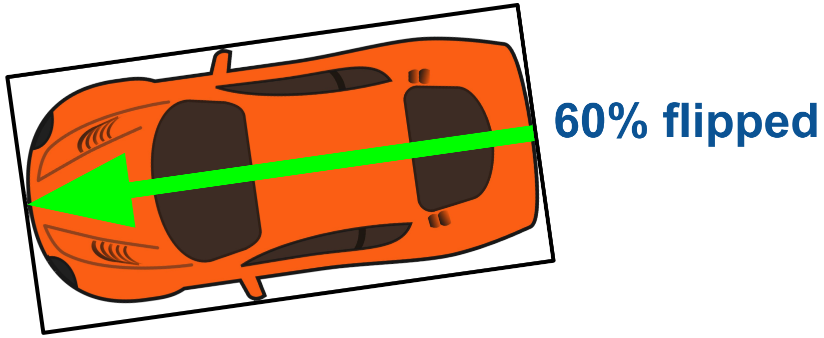
$$\mathcal{L}_{full} = \ell_1\big(\sin(\hat{\theta}) - \sin(\theta)\big) + \ell_1\big(\cos(\hat{\theta}) - \cos(\theta)\big)$$

$$\mathcal{L}_{flipped} = \ell_1\big(-\sin(\hat{\theta}) - \sin(\theta)\big) + \ell_1\big(-\cos(\hat{\theta}) - \cos(\theta)\big)$$

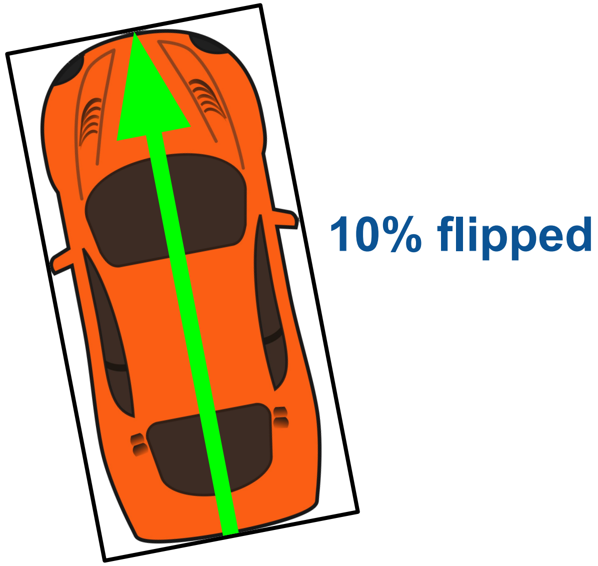
- A 180°-flipped orientation only gets penalized by the cross-entropy loss term



**Ground-truth**




**180°-flipped orientation**
Penalized by cross-entropy loss




**90°-error orientation**
High orientation loss

# Flip-aware orientation estimation

- In addition to $\sin(\theta), \cos(\theta)$, the model also outputs a flipped probability $p_f$, trained with

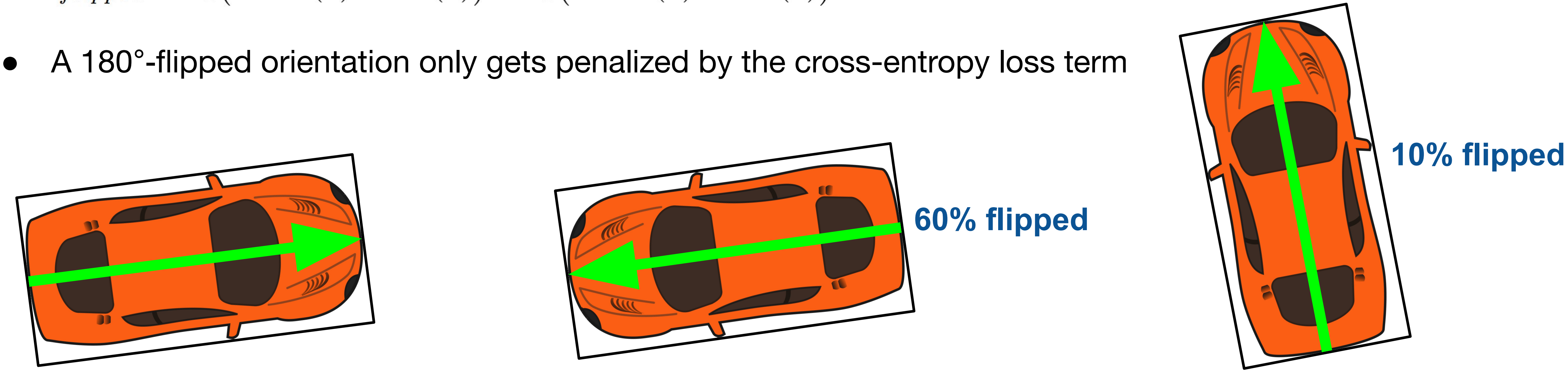
$$\mathcal{L}_{final} = \boxed{\mathcal{L}_{half} + \min(\mathcal{L}_{full}, \mathcal{L}_{flipped})} + \text{CrossEntropy}(\hat{p}_f, \mathbb{1}_{\mathcal{L}_{full} > \mathcal{L}_{flipped}})$$

where $\mathcal{L}_{half} = \ell_1\big(\sin(2\hat{\theta}) - \sin(2\theta)\big) + \ell_1\big(\cos(2\hat{\theta}) - \cos(2\theta)\big)$ ; $\sin(2\hat{\theta}) = 2\sin(\hat{\theta})\cos(\hat{\theta})$ ; $\cos(2\hat{\theta}) = \cos^2(\hat{\theta}) - \sin^2(\hat{\theta})$

$$\mathcal{L}_{full} = \ell_1\big(\sin(\hat{\theta}) - \sin(\theta)\big) + \ell_1\big(\cos(\hat{\theta}) - \cos(\theta)\big)$$

$$\mathcal{L}_{flipped} = \ell_1\big(-\sin(\hat{\theta}) - \sin(\theta)\big) + \ell_1\big(-\cos(\hat{\theta}) - \cos(\theta)\big)$$

- A 180°-flipped orientation only gets penalized by the cross-entropy loss term




**Ground-truth**

**180°-flipped orientation**
Penalized by cross-entropy loss

**90°-error orientation**
High orientation loss

# Flip-aware orientation estimation

- In addition to $\sin(\theta), \cos(\theta)$, the model also outputs a flipped probability $p_f$, trained with

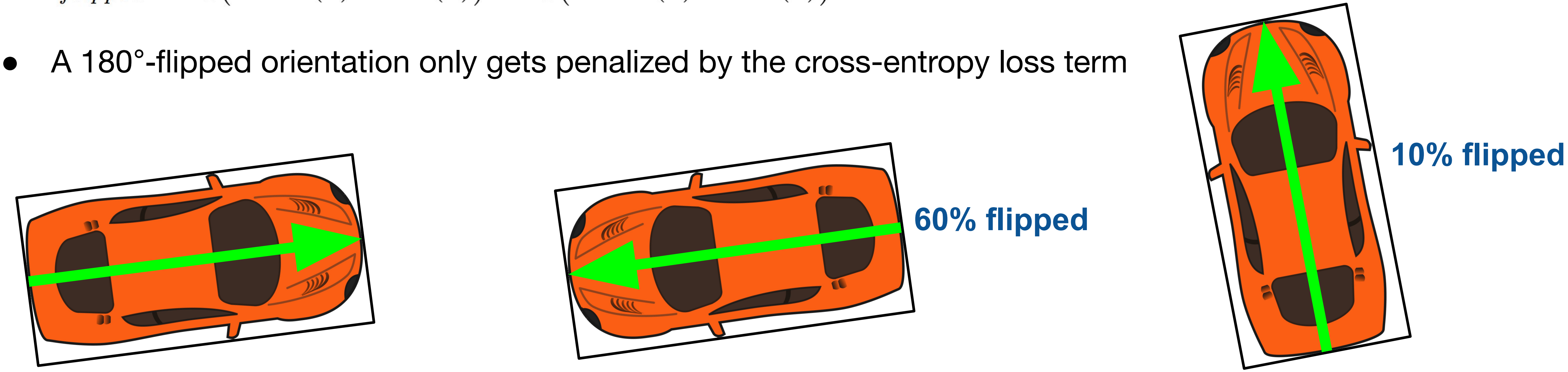
$$\mathcal{L}_{final} = \mathcal{L}_{half} + \min(\mathcal{L}_{full}, \mathcal{L}_{flipped}) + \boxed{\mathrm{CrossEntropy}(\hat{p}_f, \mathbb{1}_{\mathcal{L}_{full} > \mathcal{L}_{flipped}})}$$

where $\mathcal{L}_{half} = \ell_1\big(\sin(2\hat{\theta}) - \sin(2\theta)\big) + \ell_1\big(\cos(2\hat{\theta}) - \cos(2\theta)\big)$ ; $\sin(2\hat{\theta}) = 2\sin(\hat{\theta})\cos(\hat{\theta})$ ; $\cos(2\hat{\theta}) = \cos^2(\hat{\theta}) - \sin^2(\hat{\theta})$

$$\mathcal{L}_{full} = \ell_1\big(\sin(\hat{\theta}) - \sin(\theta)\big) + \ell_1\big(\cos(\hat{\theta}) - \cos(\theta)\big)$$

$$\mathcal{L}_{flipped} = \ell_1\big(-\sin(\hat{\theta}) - \sin(\theta)\big) + \ell_1\big(-\cos(\hat{\theta}) - \cos(\theta)\big)$$

- A 180°-flipped orientation only gets penalized by the cross-entropy loss term




**Ground-truth**

**180°-flipped orientation**
Penalized by cross-entropy loss

**90°-error orientation**
High orientation loss

# Flip-aware orientation estimation

- In addition to $\sin(\theta), \cos(\theta)$, the model also outputs a flipped probability $p_f$, trained with

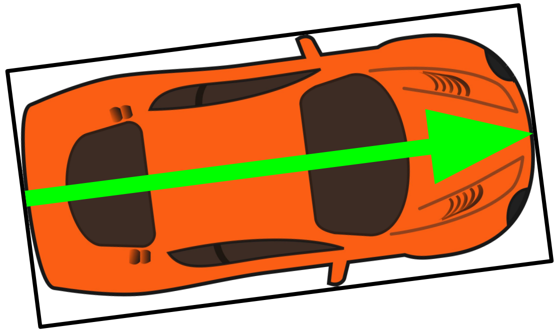
$$\mathcal{L}_{final} = \mathcal{L}_{half} + \min(\mathcal{L}_{full}, \mathcal{L}_{flipped}) + \text{CrossEntropy}(\hat{p}_f, \mathbb{1}_{\mathcal{L}_{full} > \mathcal{L}_{flipped}})$$

where $\mathcal{L}_{half} = \ell_1\big(\sin(2\hat{\theta}) - \sin(2\theta)\big) + \ell_1\big(\cos(2\hat{\theta}) - \cos(2\theta)\big)$ ; $\sin(2\hat{\theta}) = 2\sin(\hat{\theta})\cos(\hat{\theta})$ ; $\cos(2\hat{\theta}) = \cos^2(\hat{\theta}) - \sin^2(\hat{\theta})$

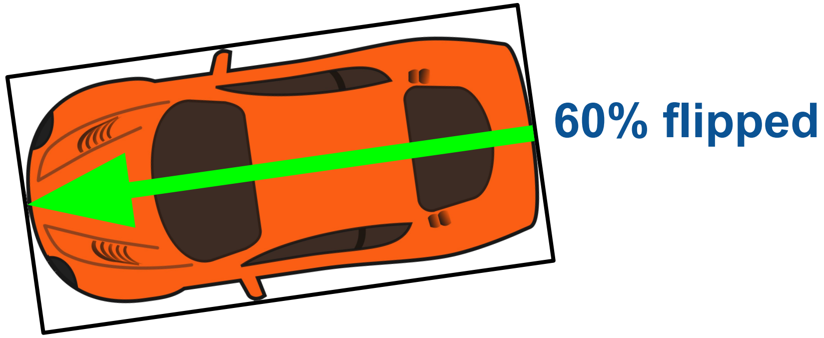
$$\mathcal{L}_{full} = \ell_1\big(\sin(\hat{\theta}) - \sin(\theta)\big) + \ell_1\big(\cos(\hat{\theta}) - \cos(\theta)\big)$$

$$\mathcal{L}_{flipped} = \ell_1\big(-\sin(\hat{\theta}) - \sin(\theta)\big) + \ell_1\big(-\cos(\hat{\theta}) - \cos(\theta)\big)$$

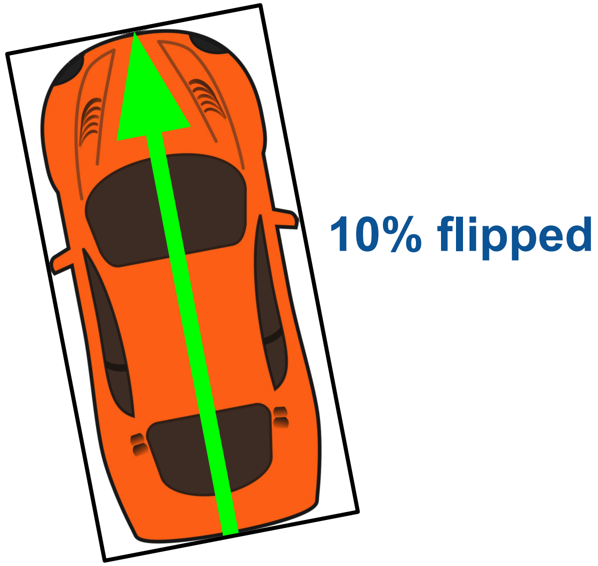
- A 180°-flipped orientation only gets penalized by the cross-entropy loss term



**Ground-truth**




**180°-flipped orientation**
Penalized by cross-entropy loss




**90°-error orientation**
High orientation loss

# Highlights of evaluation results on nuScenes

| Method | Average Precision$_{0.7}$ ↑ | Average Orientation Similarity$_{0.7}$ ↑ |
|---|---|---|
| Naive baseline | $57.1 \pm 0.5$ | $55.1 \pm 0.5$ |
| MultiBin baseline | $58.0 \pm 0.5$ | $55.5 \pm 0.3$ |
| Flip-aware | **60.7** $\pm$ **0.2** | **57.9** $\pm$ **0.4** |

- Applied our and baseline methods on a state-of-the-art joint detection prediction model
- Average precision (AP) for detection and average orientation similarity (AOS) for orientation
- Flip-aware achieved significant improvements over the baseline methods
- Please read our paper for more details