Rethinking of Radar's Role: A Camera-Radar Dataset and Systematic Annotator via Coordinate Alignment

Yizhou Wang[1], Gaoang Wang[2], Hung-Min Hsu[1], Hui Liu[1,3], Jenq-Neng Hwang[1]

[1] University of Washington, [2] Zhejiang University, [3] Silkwave Holdings Limited

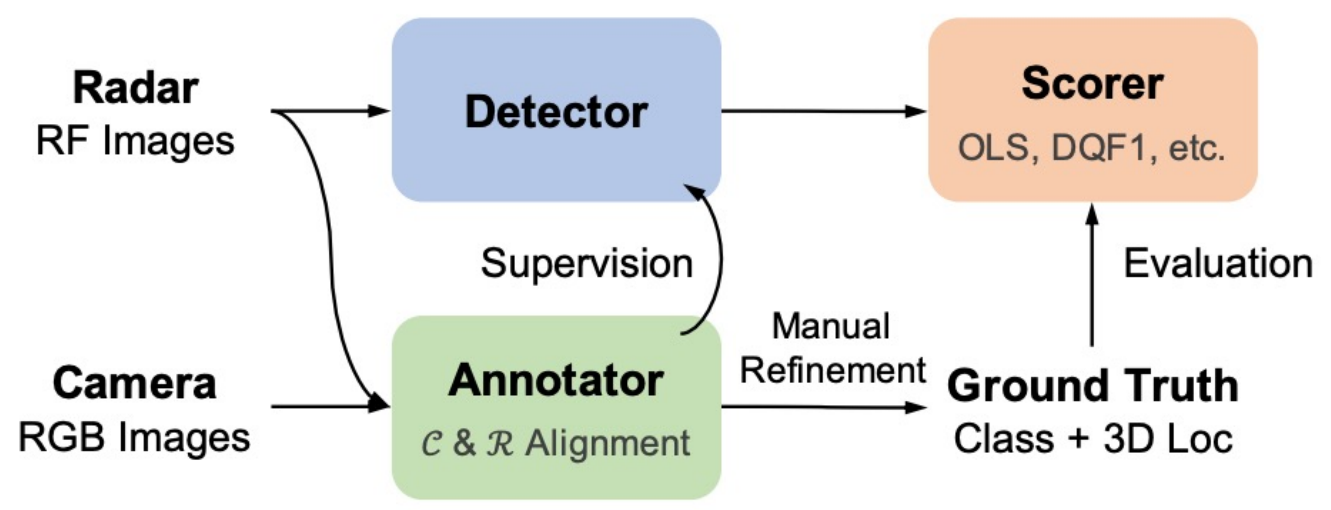
CVPR VIRTUAL JUNE 19-25

Motivation

- ***Traditional Radar's Role for Autonomous Driving***
- Obstacle ranging and speed estimation.
- All-weather robust sensing.
- Supplementary sensor for camera, LiDAR, etc.

- ***Rethinking of Radar's Role***
- Use as a primary sensor independently.
- Achieve more complicated tasks, e.g., object detection and tracking.
- It is very important when other sensors are not reliable in **adverse** driving scenarios.

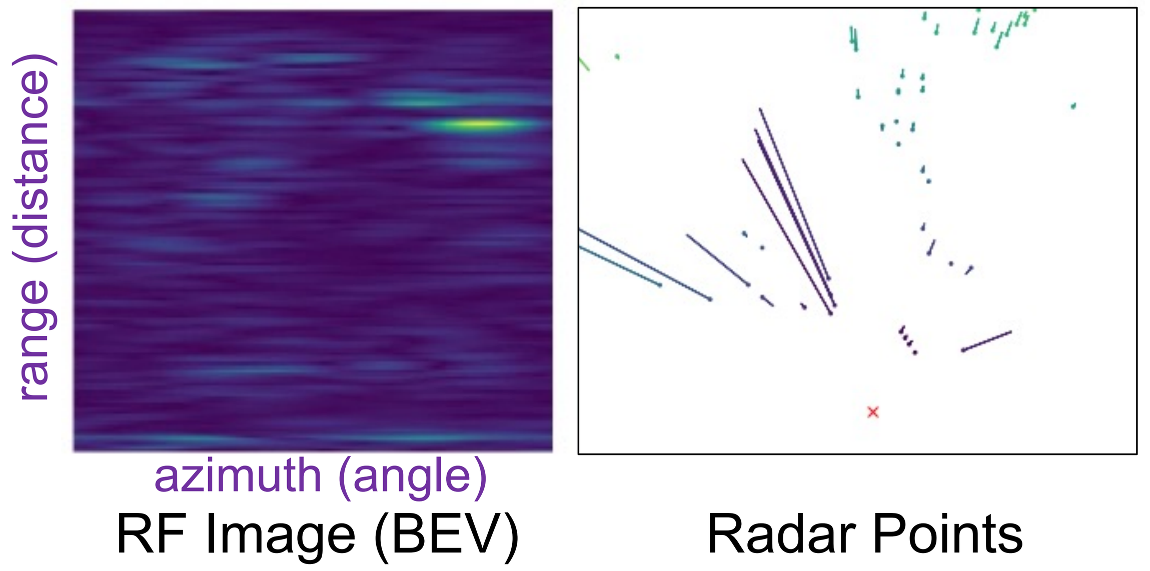
- ***Radar Object Detection (ROD)***



- **Detector**: uses radar data as the only input to do object detection.
- **Annotator**: takes both camera and radar data to generate annotations by detection alignment between camera and radar.
- **Scorer**: compare the detection with ground truth using a series of evaluation metrics.

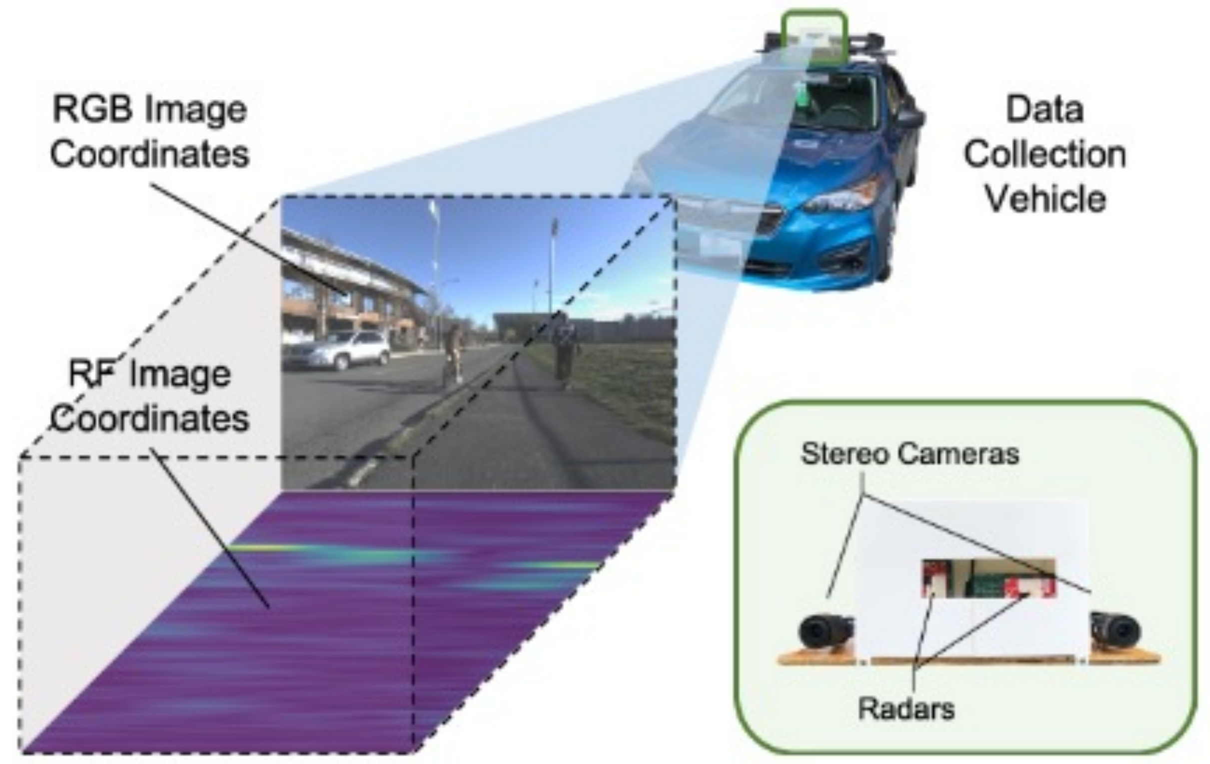
Dataset

- ***Radar Data Representations***



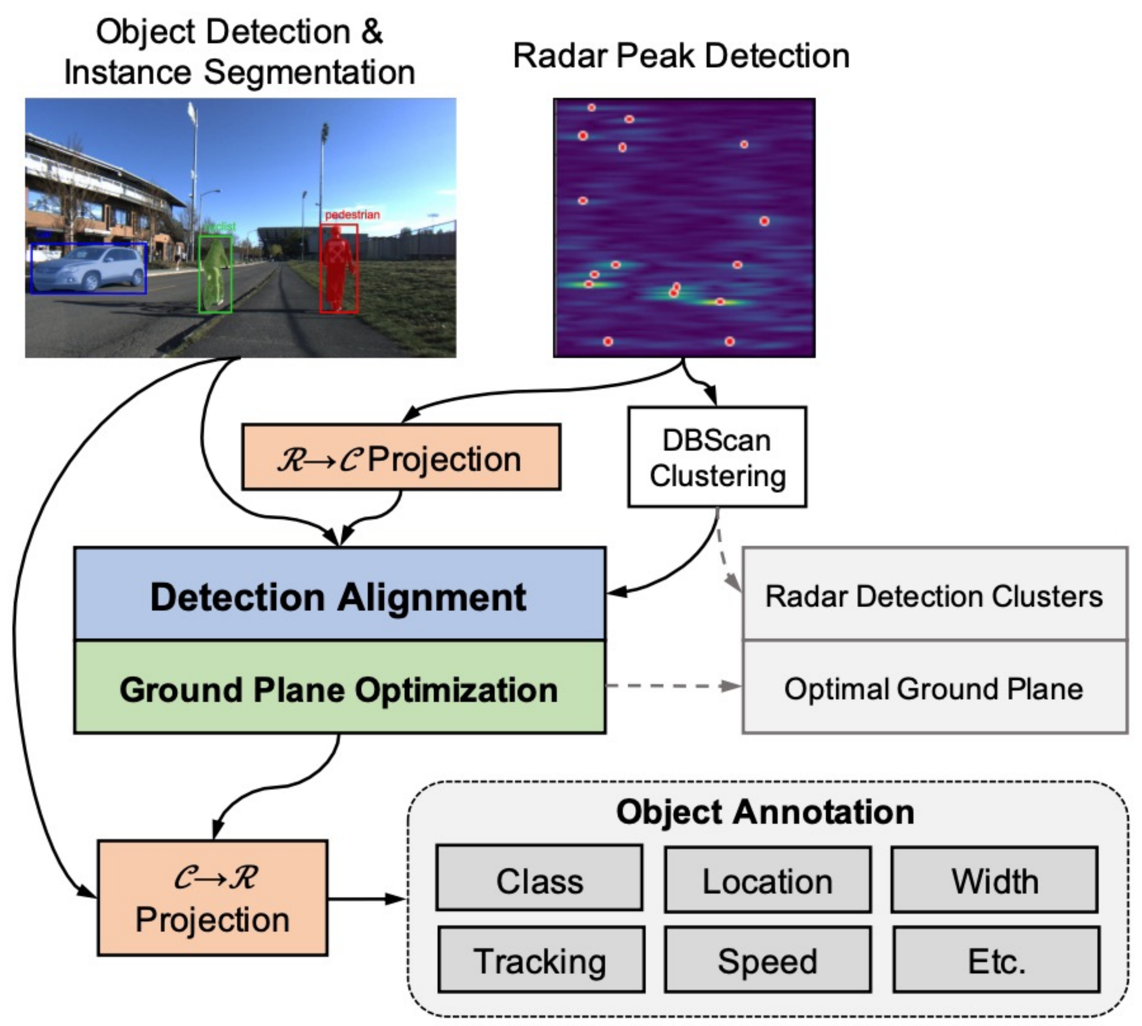
- **Radio frequency (RF) image** (Ours):
- Strength: **Rich** information including object location, shape, surface texture, speed, etc.
- Weakness: Hard to extract features.

- ***CRUW Dataset***
- **Scale**: 3.5 hours, 30 FPS, **RF images**
- **Sensors**: 1) A pair of stereo cameras. 2) Two 77GHz FMCW radar antenna arrays.
- **Scenarios**: Parking lot (PL), Campus road (CR), City street (CS), Highway (HW)



Annotator

- ***Radar Object Annotation System***



- **Initialization:**
- Object detection on RGB images
- Peak detection on RF images
- Initial clustering: DBScan
- **Bilateral Coordinate Projection:**
- Project detection between RGB pixel coordinates and radar range-azimuth coordinates through ground plane.
- **Detection Alignment and Optimization:**
- Detection alignment cost

$$\ell_i = \lambda_i \left(h_i - h_i^{mask}\right)^2 + (1-\lambda_i)\left(h_i - h_i^{bbox}\right)^2$$

- Ground plane optimization

$$\min_{\varphi,\gamma} \sum_t \sum_{i=1}^{n_{CFAR}} \left(v_{i,t} - v_{i,t}^{mask}\right)^2$$

Scorer

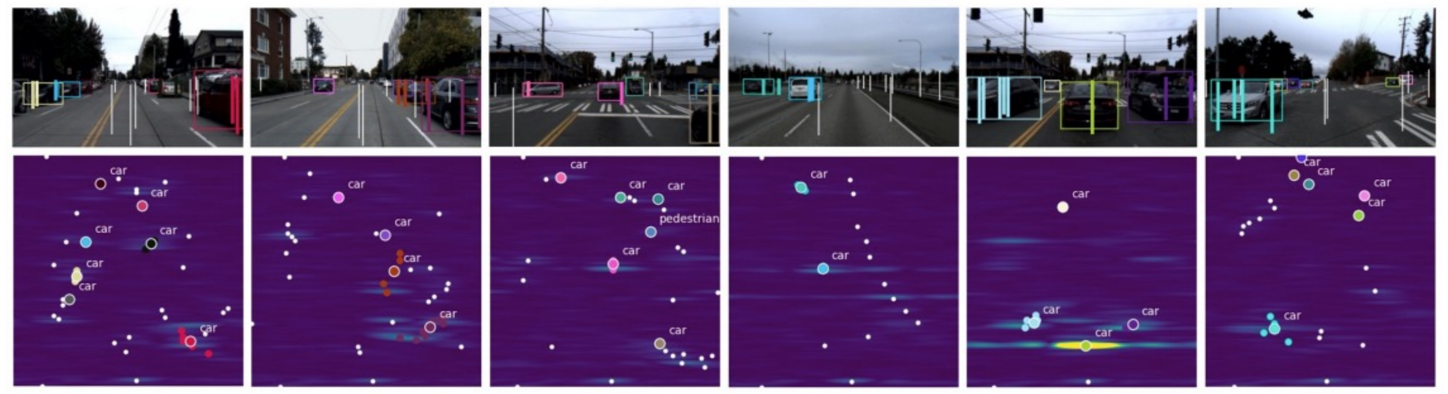
- ***Point-based Similarity***

$$\mathrm{OLS}(i,j) = \exp\left\{\frac{-d_{ij}^2}{2(s_j\kappa_{cls})^2}\right\}$$

d: distance between two points.
s: distance from target to sensor.
κ: per-class constant for error tolerance.

- ***Evaluation Metrics***
- Mean absolute error (MAE)
- Average precision (AP)
- Average recall (AR)
- Detection Quality F1 (DQF1) Score:

$$\mathrm{DQF1} = \frac{2}{n_{det} + n_{gt}} \sum_{j=1}^{n_{gt}} \sum_{i=1}^{n_{det}} \delta_{i,j} \cdot \mathrm{OLS}(i,j)$$



Method	Scenario	MAE	Precision	Recall	**AP**	**AR**	**DQF1**
RODNet (Vanilla) [33]	Overall	0.31 (±0.26)	95.90%	78.03%	74.29%	77.85%	81.02%
	Parking Lot	0.26 (±0.19)	98.29%	87.76%	85.33%	86.76%	89.33%
	Campus Road	0.42 (±0.30)	89.49%	53.02%	42.67%	49.03%	56.03%
	City Street	0.48 (±0.39)	88.88%	73.42%	59.79%	67.23%	71.15%
RODNet (HG) [33]	Overall	0.31 (±0.23)	96.02%	88.56%	83.76%	85.62%	86.64%
	Parking Lot	0.26 (±0.16)	98.26%	96.94%	93.60%	94.98%	93.63%
	Campus Road	0.40 (±0.26)	92.16%	68.76%	50.34%	57.23%	70.28%
	City Street	0.48 (±0.39)	91.53%	81.27%	64.54%	70.47%	75.55%
RODNet (Full) [34]	Overall	0.31 (±0.25)	95.93%	88.86%	85.98%	87.86%	87.82%
	Parking Lot	0.27 (±0.21)	98.49%	97.98%	95.79%	96.85%	94.62%
	Campus Road	0.36 (±0.26)	92.08%	69.40%	57.06%	62.08%	73.62%
	City Street	0.49 (±0.37)	91.59%	76.37%	62.83%	70.41%	74.65%