



Echoes Beyond Points: Unleashing the Power of Raw Radar Data in Multi-modality Fusion

王峰
算法研究员
图森未来



传统4D radar信号处理流程

距离：

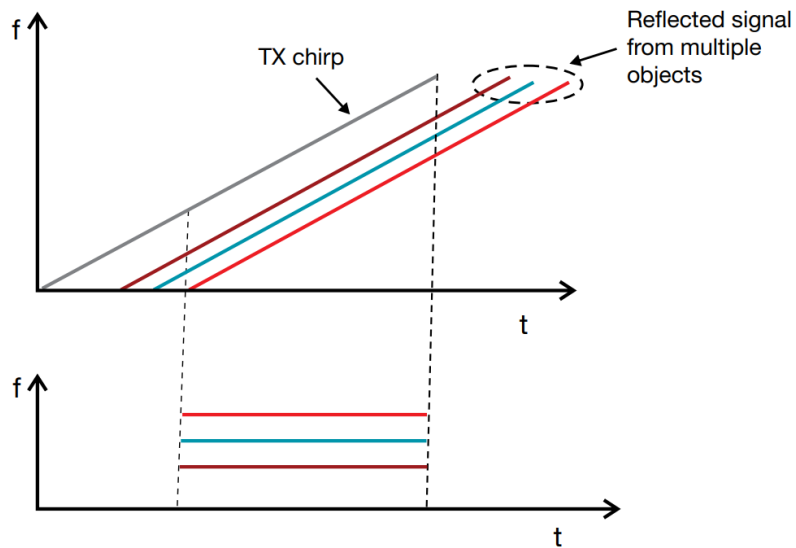


Figure 5. Multiple IF tones for multiple-object detection.

速度：

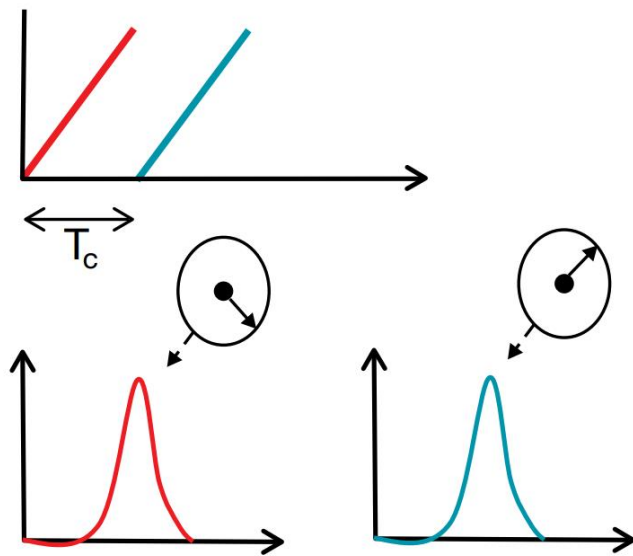
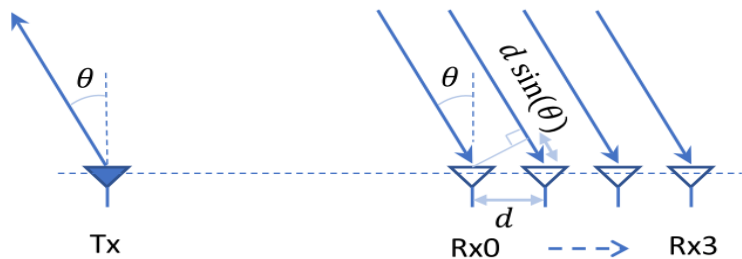


Figure 6. Two-chirp velocity measurement.

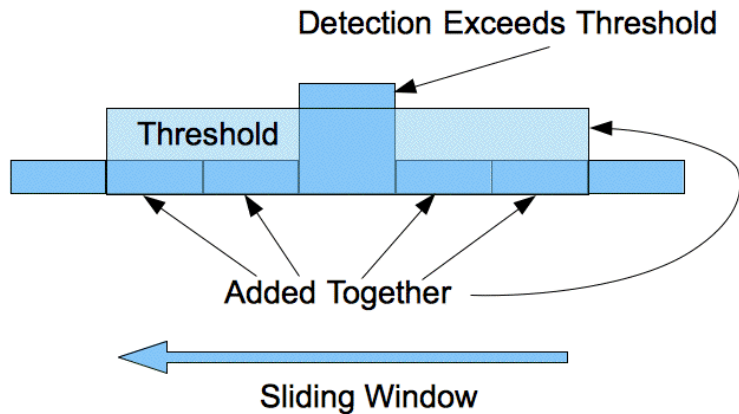


传统4D radar信号处理流程

角度：

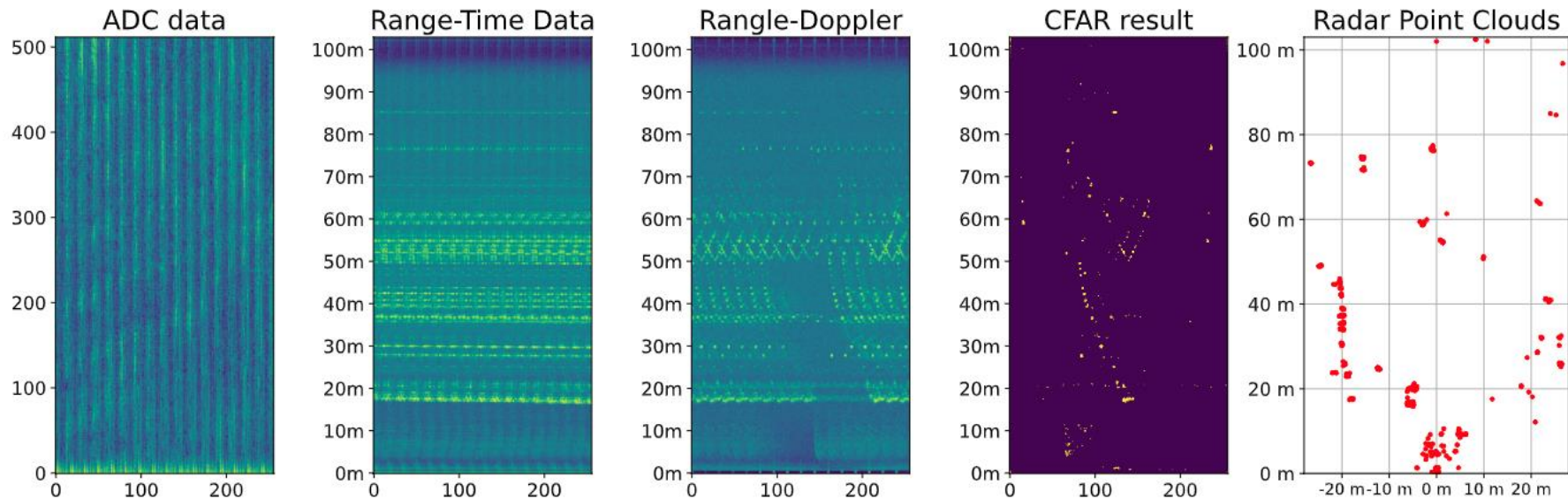


CFAR选点：





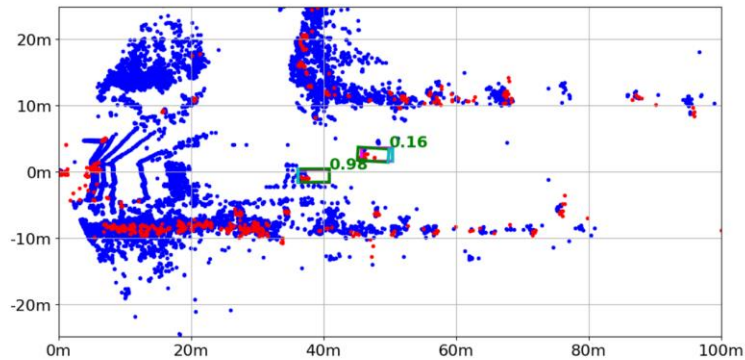
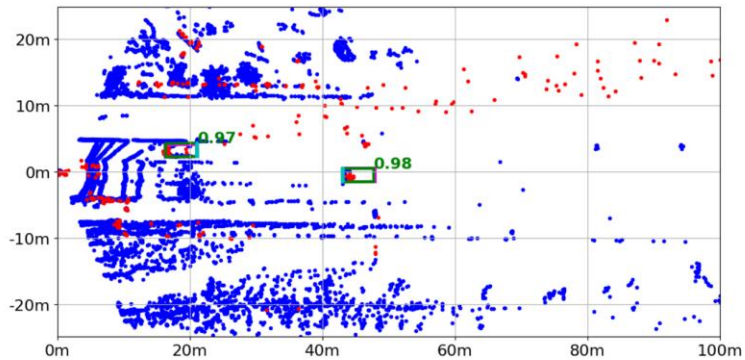
4D radar原始数据信号处理流程



Radar原始数据的不同转换形式

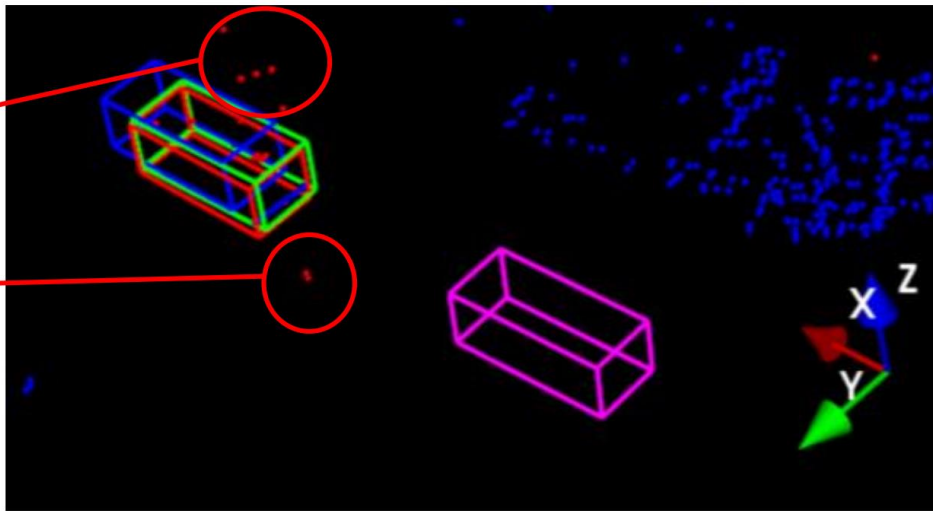
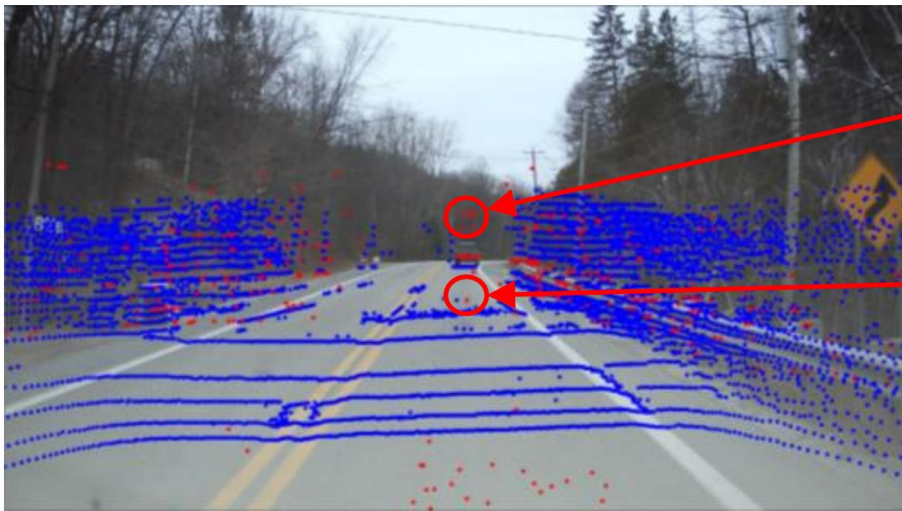
传统信号处理方案损失了大量信息！

纯Radar目标检测的缺陷 – 角度分辨率不足





纯视觉目标检测的缺陷 – 缺乏深度信息



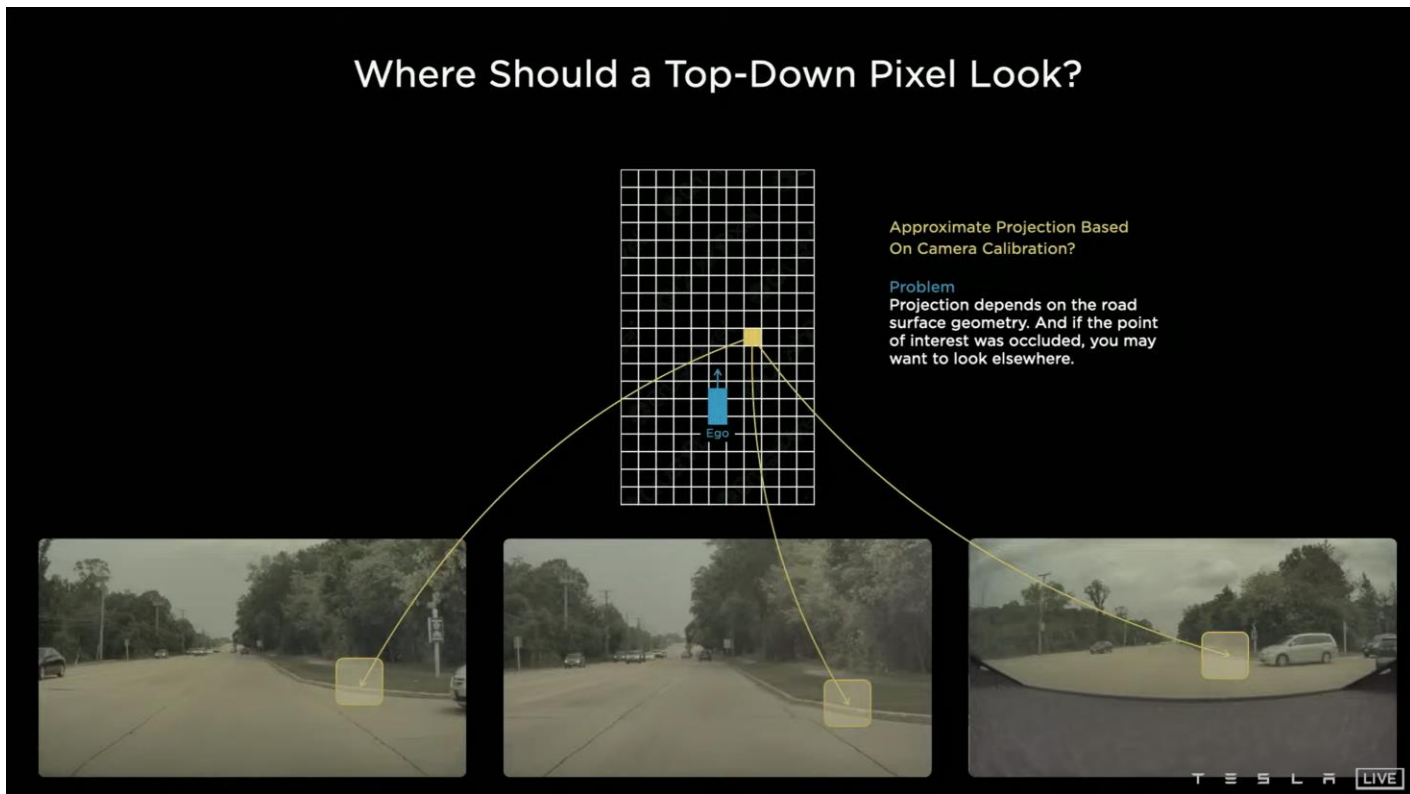
蓝点：LiDAR点 红点：4D Radar点

红框：真值 紫框：纯图像BEV检测 蓝框：图像与radar点融合 绿框：图像与radar原始数据融合

深度不准确是纯视觉目标检测最常见的错误。

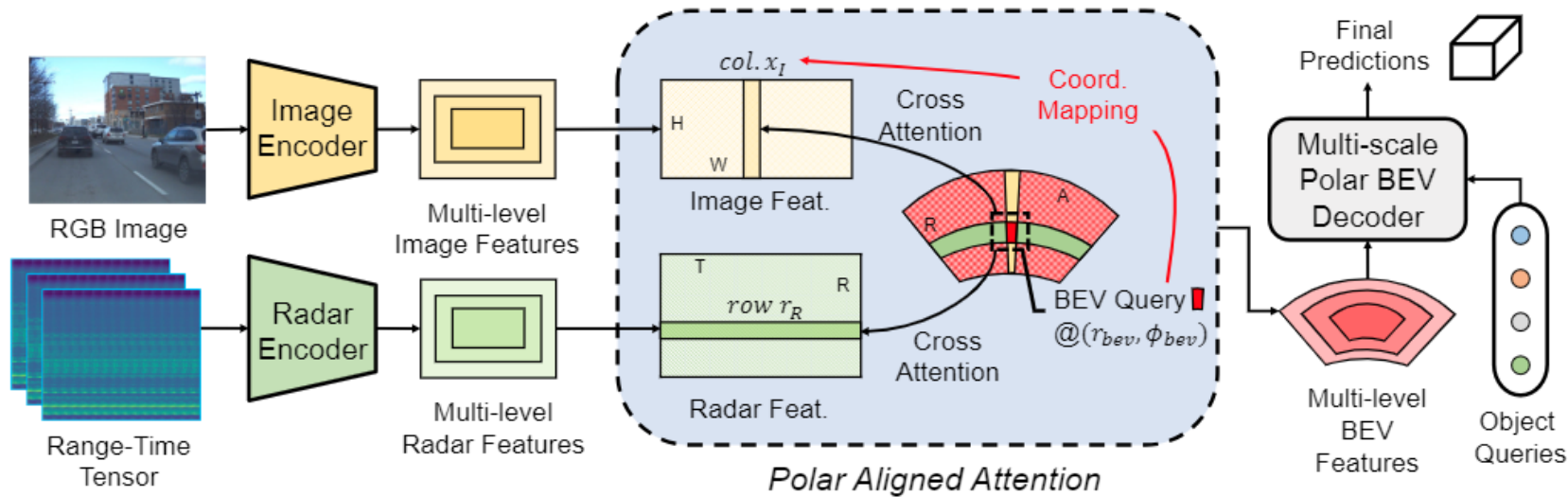


BEV视角下的纯视觉目标检测





4D原始数据融合带来巨大性能提升



将Radar原始数据和图像在BEV空间中融合

图像缺乏深度、Radar角度分辨率不足。在BEV框架下，二者恰好互补！



评测数据集

- RADlal: 图像+4D Radar原始ADC数据, 只有物体中心的标注。
- RADlal+: 增加了旋转框标注。
- KRadar: 图像+4D Radar RA map数据, 有旋转框标注。数据集难以下载。

| Dataset | Data | Other Sensors | Scenarios | Annotations | Classes | Size |
|--------------------|-----------|---------------|-----------|-------------|---------|------|
| Astyx [28] | PC | CL | SH | 3D | 7 | 500 |
| View-of-Delft [37] | PC | C_s LO | U | 3D,T | 13 | 8693 |
| RADlal [39] | ADC,RA,PC | CLO | USH | P_o ,M | 1 | 8252 |
| TJ4DRadSet [56] | PC | CLO | U | 3D,T | 5 | 7757 |
| Radatron [26] | ADC,RA | C_s | U | BEV | 1 | 16K |
| KRadar [35] | RA, PC | CLO | USH | 3D, T | 5 | 17K |



4D原始数据融合带来巨大性能提升

和图像融合后，在目标检测任务中接近LiDAR性能，大幅超越原始Radar效果。

相比于点云，原始数据在远距离的表现更好。

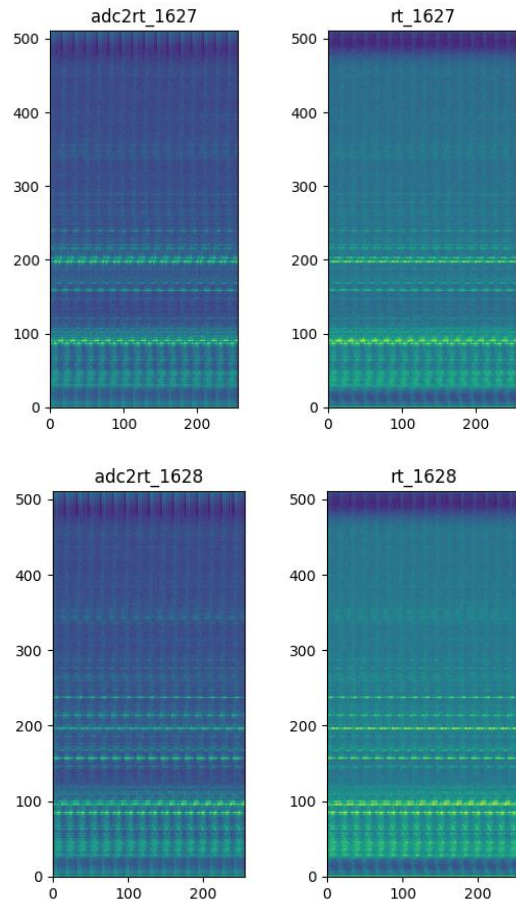
| Modality | | | | | BEV AP@0.7(%)↑ | | | LET-BEV-AP@0.7(%)↑ | | |
|----------|----|----|----|---|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|
| C | RT | RA | RP | L | Overall | 0 - 50m | 50 - 100m | Overall | 0 - 50m | 50 - 100m |
| ✓ | | | | | 10.07±1.10 | 21.68±3.11 | 2.21±0.59 | 56.92±3.91 | 78.98±1.41 | 36.20±5.81 |
| | ✓ | | | | 48.26±0.92 | 62.33±4.00 | 45.05±0.98 | 55.04±1.25 | 68.56±4.11 | 52.30±1.82 |
| | | ✓ | | | 50.67±1.03 | 65.05±1.59 | 42.98±0.51 | 58.53±0.13 | 72.95±0.78 | 51.41±0.54 |
| | | | ✓ | | 53.55±0.70 | 66.19±1.72 | 47.41±1.06 | 62.59±1.14 | 76.45±2.06 | 56.73±1.33 |
| | | | | ✓ | 84.47±0.15 | 86.92±0.61 | 91.91±0.13 | 86.07±0.44 | 88.55±0.79 | 93.17±0.10 |
| ✓ | ✓ | | | | 84.92±0.98 | 87.56±1.58 | 91.06±1.18 | 88.86±0.62 | 92.81±1.37 | 94.81±0.52 |
| ✓ | | ✓ | | | 84.77±0.65 | 87.93±0.60 | 91.48±0.28 | 89.54±0.54 | 93.83±0.56 | 94.87±0.58 |
| ✓ | | | ✓ | | 82.35±0.93 | 86.97±1.01 | 86.39±0.74 | 88.35±0.78 | 93.07±0.80 | 92.77±0.69 |
| ✓ | | | | ✓ | 86.35±1.15 | 88.81±1.40 | 94.25±1.68 | 88.44±0.79 | 91.63±1.34 | 95.31±1.28 |

| Training Dataset | Vehicle AP/Conf@0.3 | AP_BEV@0.3 | AP_3D@0.3 | AP_BEV@0.5 | AP_3D@0.5 | AP_BEV@0.7 | AP_3D@0.7 |
|--------------------|---------------------|--------------|---------------|---------------|---------------|---------------|---------------|
| KRadar | RTHN-Pretrained | 58.04 | 49.65 | 42.60 | 17.87 | 10.69 | 0.45 |
| KRadar-20-trainval | RTHN-Scratch | 61.38 | 53.05 | 46.47 | 17.98 | 10.47 | 3.03 |
| KRadar-20-trainval | EchoFusion-ra+img | 69.95(+8.63) | 68.35(+15.30) | 57.28(+10.81) | 43.87(+25.89) | 33.07(+22.60) | 14.00(+10.97) |
| KRadar-20-train | EchoFusion-rpcd | 55.74 | 53.40 | 43.94 | 29.21 | 20.56 | 2.90 |
| KRadar-20-train | EchoFusion-ra | 54.45 | 51.75 | 42.28 | 24.65 | 17.97 | 3.85 |
| KRadar-20-train | EchoFusion-rpcd+img | 66.46 | 58.62 | 54.17 | 34.67 | 26.39 | 3.96 |
| KRadar-20-train | EchoFusion-ra+img | 68.70 | 66.68 | 55.90 | 34.33 | 29.65 | 5.34 |



原始ADC信号

| RADIAL-Test- BEV@0.7 | <u>mAP</u> (%)↑ | <u>mAP</u> (%) ↑ [0, 50] | <u>mAP</u> (%) ↑ [50, 100] | let-mAP(%)↑ | let-mAP(%) ↑ [0, 50] | let-mAP(%) ↑ [50, 100] |
|--|-----------------|-----------------------------|-------------------------------|-------------|-------------------------|---------------------------|
| rt_radarformer_r50_da, 21e | 48.03 | 58.96 | 46.43 | 55.19 | 66.13 | 54.87 |
| adc_fourier_range_embed_0.1lr_MP, 14e | 47.68 | 58.29 | 46.74 | 54.42 | 65.72 | 53.68 |
| rt_img_radarformer_r50_da_small_head_3layer, 16e | 83.07 | 89.42 | 92.65 | 86.93 | 90.79 | 93.22 |
| adc_img_radarformer_r50_da_MP_3layer, 16e | 82.86 | 87.26 | 88.70 | 87.38 | 92.57 | 93.49 |

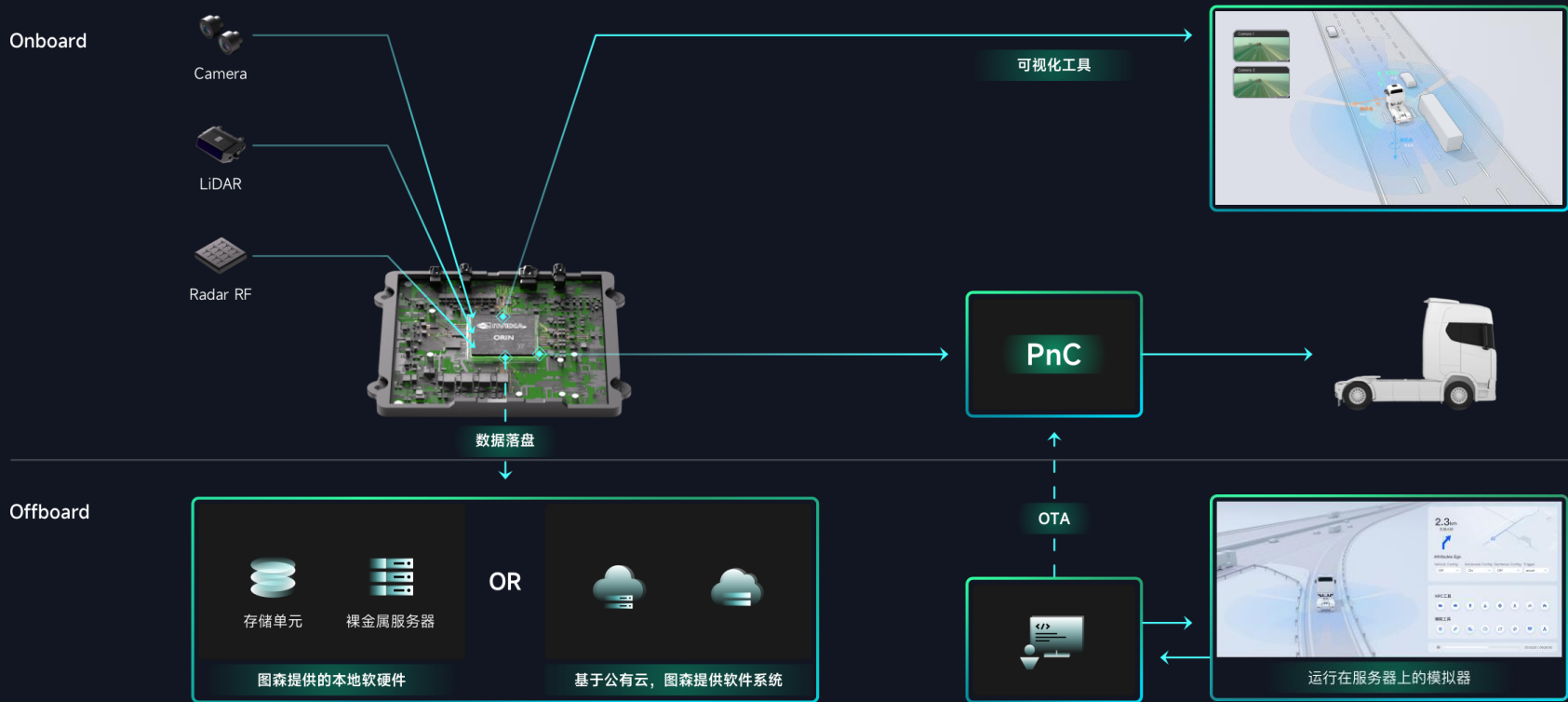


即使不做Range-FFT，网络依然可以学到类似的变换。



小结

- 原始数据相比于点云，没有信息损失，在远距离检测中优势明显。
- FFT等效于一个linear layer，网络可以隐式学习到传统雷达信号处理的过程。
- 图像缺乏深度，Radar角度分辨率较低，两者刚好互补。
- 原始数据的数据量较大，需要很强的系统设计和集成能力。





全栈自研智能驾驶解决方案

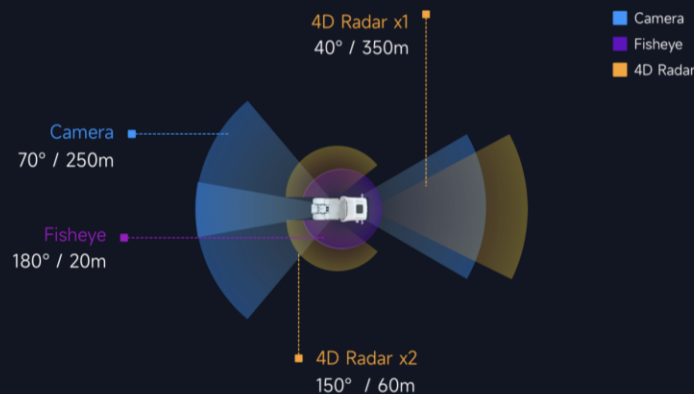
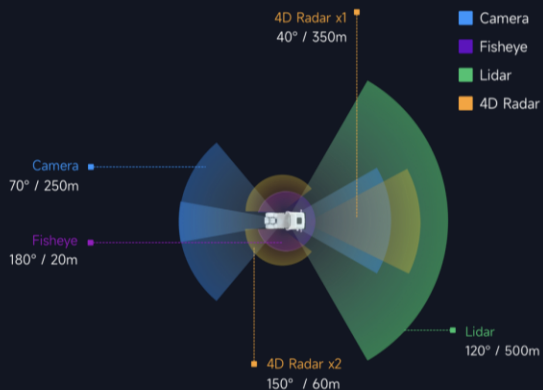
“节油领航”系列产品

搭载在自研 TS-BOX 上

集成传感器输入、高性能中间件、车辆控制单元(VCU)
和自动驾驶应用软件

采用自研自动驾驶算法，通过 ETA (预计到达时间)计算
实现最高效的节油

节油、安全，满足高速点到点领航



“节油领航” 传感器布局（可选配前向LiDAR）

已获两家OEM量产项目定点，24年底SOP



域控集中式4D Radar性能更好

感知方案总包，减少OEM烦恼

工具链友好，研发更高效

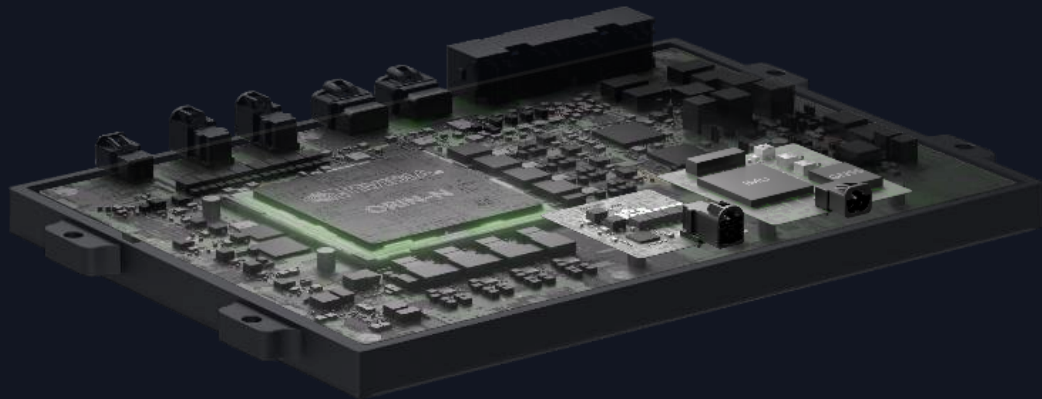
整套方案成本更低: GNSS / INS系统成本更低

商用车和乘用车都支持

成本降低



25%





谢谢大家！

文章链接: <https://arxiv.org/abs/2307.16532>

开源代码: <https://github.com/tusen-ai/EchoFusion>