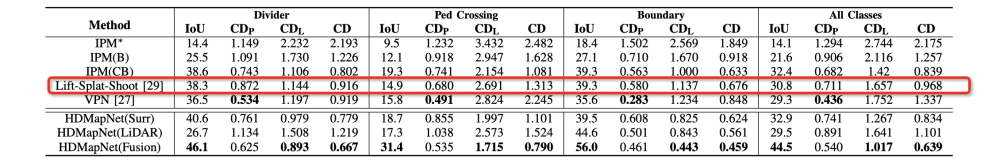
# BMBEV：+Seg

### 1、HDMapNet模型复现

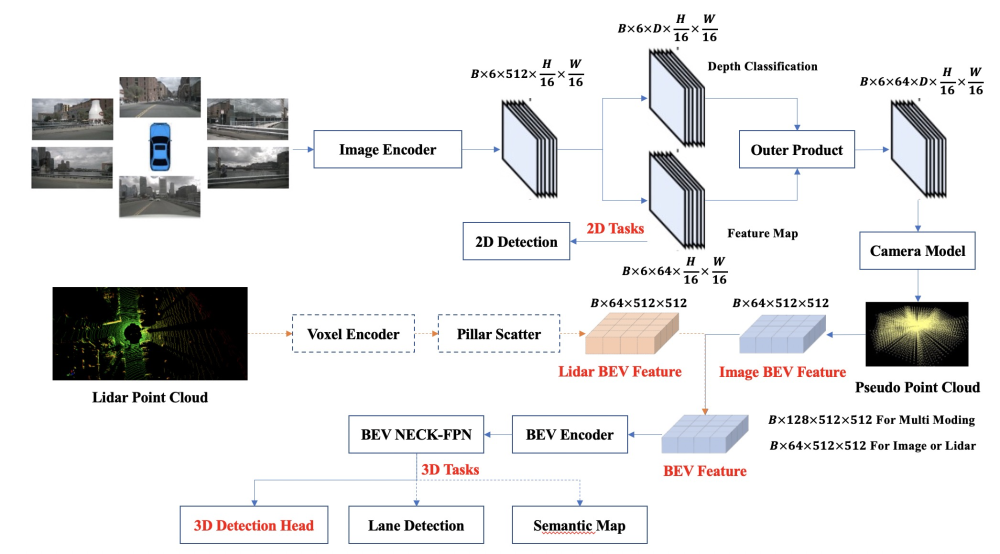
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Divider | Ped Crossing | Boundary | All Classes |  |
| HDMapNet(Paper-Surr) | 40.6 | 18.7 | 39.5 | 32.9 |  |
| HDMapNet(Paper-Fusion) | 46.1 | 31.4 | 56.0 | 44.5 |  |
| HDMapNet(Experiment-Surr) | 38.5 | 19.3 | 38.1 | 32.0 |  |
| LSS-BM(Experiment-Surr) | 38.3 | 17.9 | 40.2 | 32.1 | 23 epoch |
| LSS-BM(Experiment-Fusion) | 49.0 | 34.4 | 58.8 | 47.4 | 28 epoch |

#### 对比论文中结果

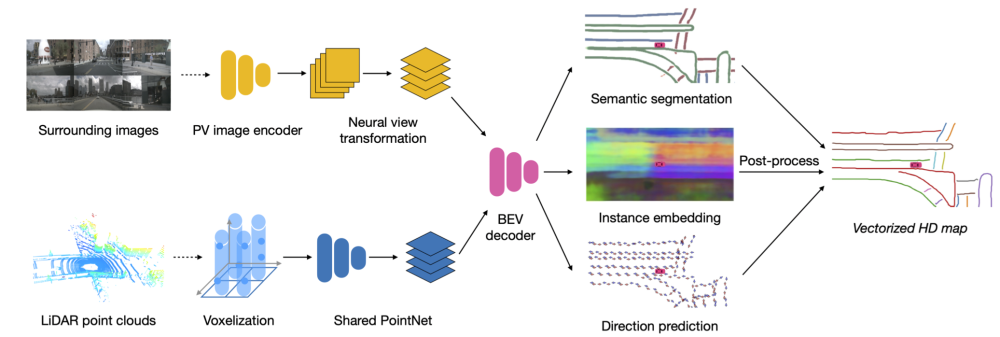


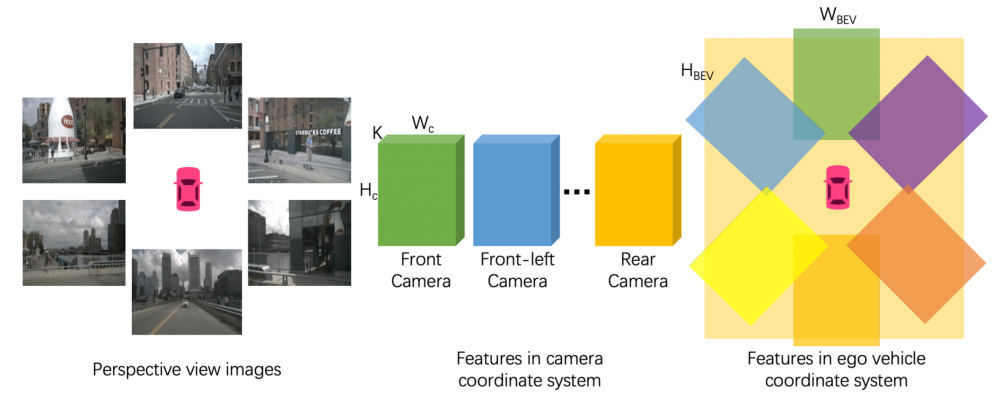
#### 框架对比

**LSS-BM:**



**HDMapNet:**





**对比分析**

* 前提设定

* + 为了对比设定，除了**view transformation**之外，image backbone、bev encoder、bev neck-fpn、pointpilars全部基于同样的模型参数配置

* + 从HDMapNet的代码来看，使用的方法是基于IPM的方式来实现BEV视角生成，如上图所示

* + 现有架构是基于LSS的方式生成BEV视角

* 结论分析

* + 直接基于开源的HDMapNet的代码进行训练，不做参数修改(代码未进行任何修改)，最终训练出来的性能average iou=32.0，与论文提供的32.9存在些许差距，原因可能是由于不同精度的训练设备导致的(无法验证)，总体上相差不多

* + 基于LSS-BM框架+HDMapNet的数据最终训练的结果average iou=**32.1**，与HDMapNet基于环视相机训练出的结果(重新训练**32.0**)持平，比论文中提供的复现的Lift-Split-Shoot的结果(**30.8**)要高**1.3%**

* + HDMapNet代码中实现BEV视角变换是基于IPM来实现(基于相机内外参直接变换)，前提假设是地面是平的假设，理论上的车道线等检测结果要由于基于深度估计的模型，但是对于可行驶区域或者边界的检测结果要比基于LSS的方法差，原因分析在于边界分界线并没有明确的分割线作为标志，多目相机融合之后，边界处的情况复杂加大了预测难度，因此对于边界boundary预测HDMapNet要比LSS-BM低了**2.1%**

* + 从实验结果来看，基于LSS的相机BEV融合多模态特征要比基于IPM生成BEV视角融合多模态特征性能上要好，由于基于IPM的变换方式，在进行训练时很难对图像层面进行数据增强操作，如对图像进行随机旋转后，直接基于IPM变换至BEV后会导致GT同样产生旋转，会加大模型学习难度

* + 总结来看，基于LSS-BM的框架进行语义分割等任务效果不输于HDMapNet(IPM)的效果，但是在多模态扩展上效果要比HDMapNet要好很多，高出**2.9%**

**语义分割结果:**

