# BMBEV：+Temporal Feat Fusion

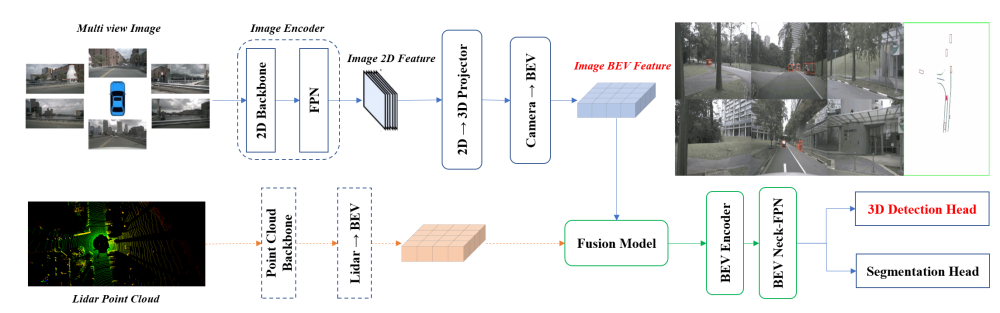
## BM-BEV框架

* **现有基于BEV的感知架构如下所示，现阶段支持:**

* + **基于LSS的Camera -> BEV的变换方式**

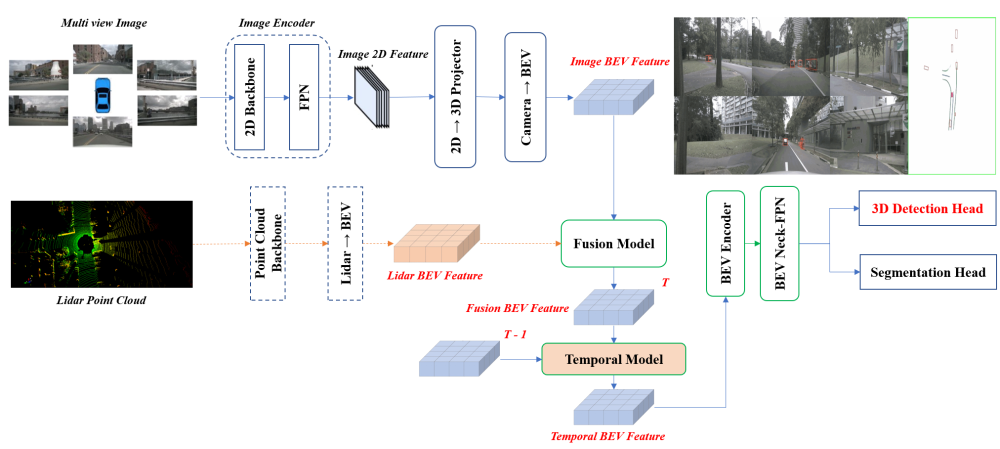
* + **Camera Only & Camera Lidar Fusion**

* + **3D Detection & Segmentation**



## 时序信息融合

* **更新后的架构如图所示:**

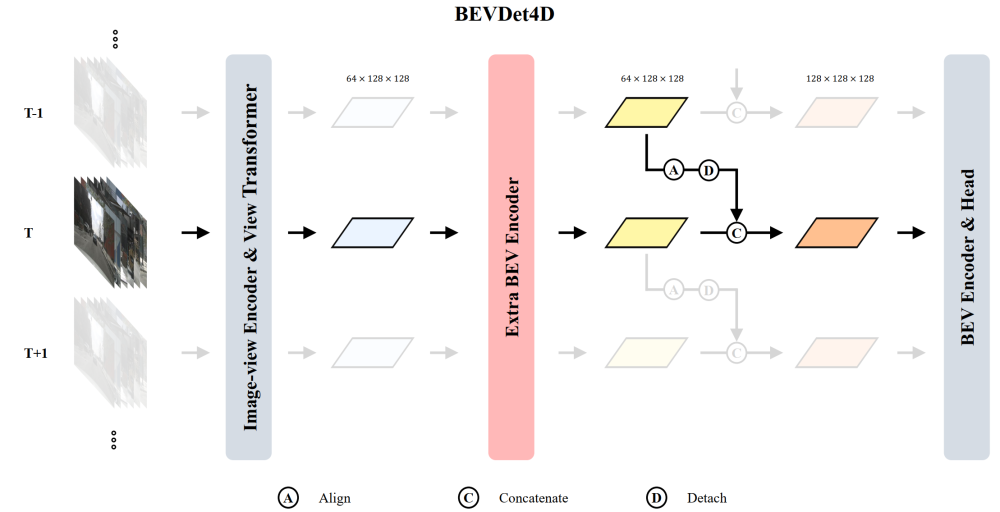


### 时序模块

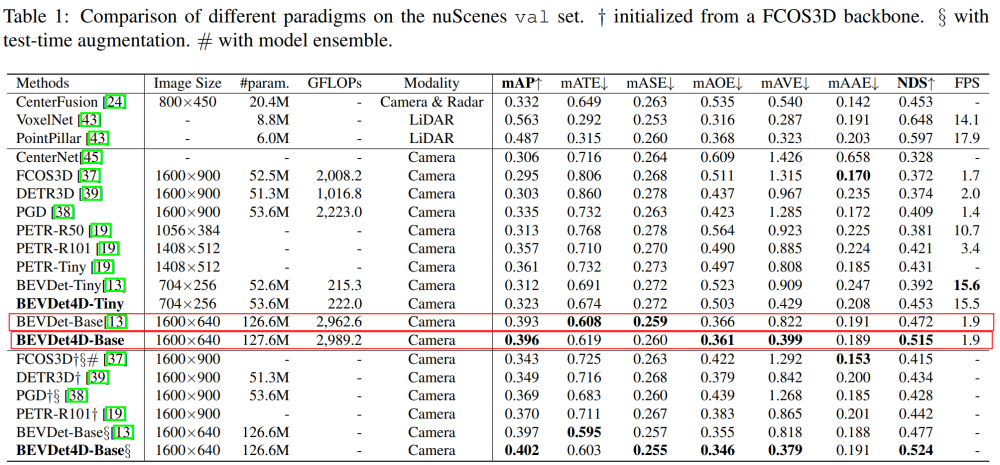
* **论文调研，目前融合了时序**

* + **BEVDet4D: 直接对齐不同时刻的特征图再进行拼接，从最后的指标来看提升不是很明显。**

* + - **BEVDet4D时序融合部分架构**

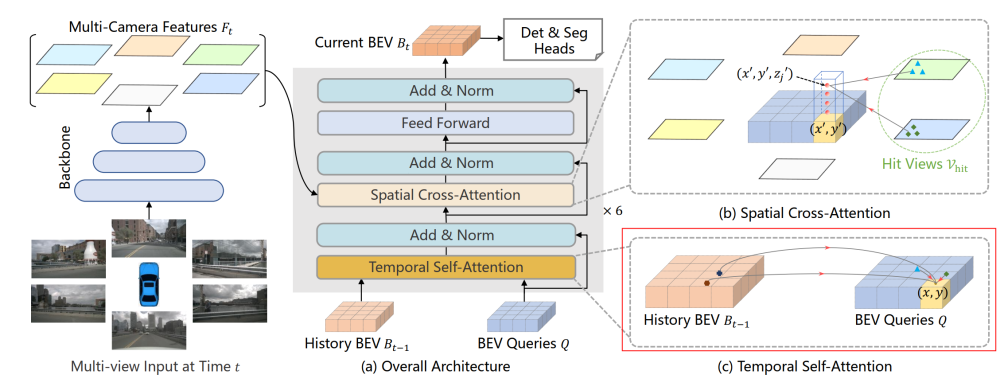


* + - **关键指标**

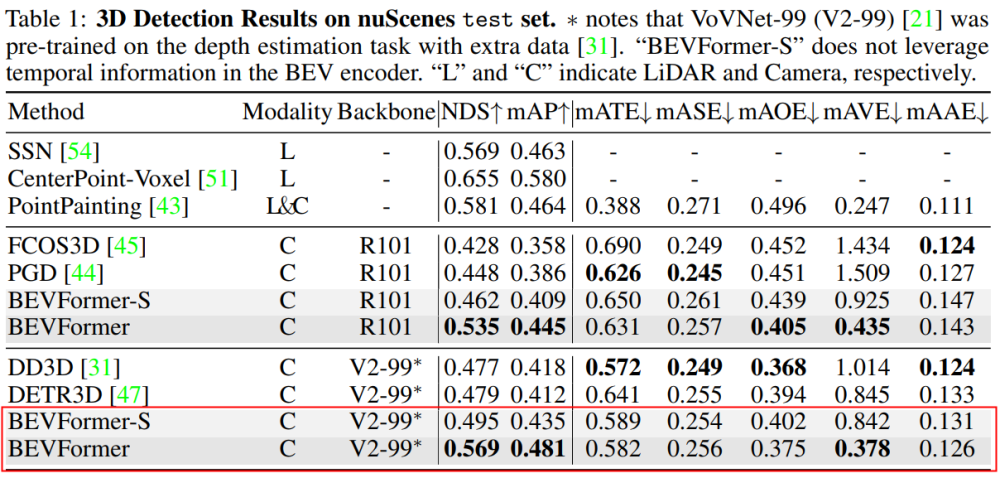


* + **BEVFormer-基于Transformer的方法来融合时序信息，存在模型过重的问题，Batch\_Size=1的情况下，显存占用用到了28G，且目前只支持Batch\_Size=1来训练，Batch\_size>1会报错**

* + - **模型架构**

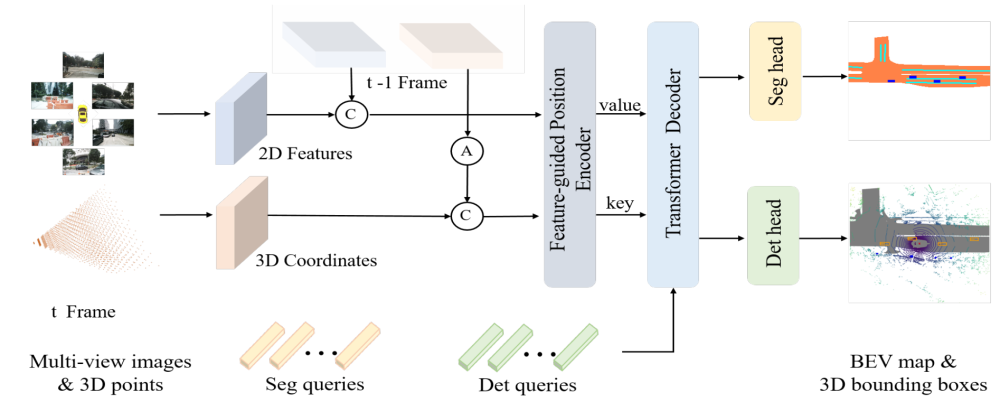


* + - **关键指标-提升还是很明显的**



* + **PETR-V2: 比较特殊的融合方式**

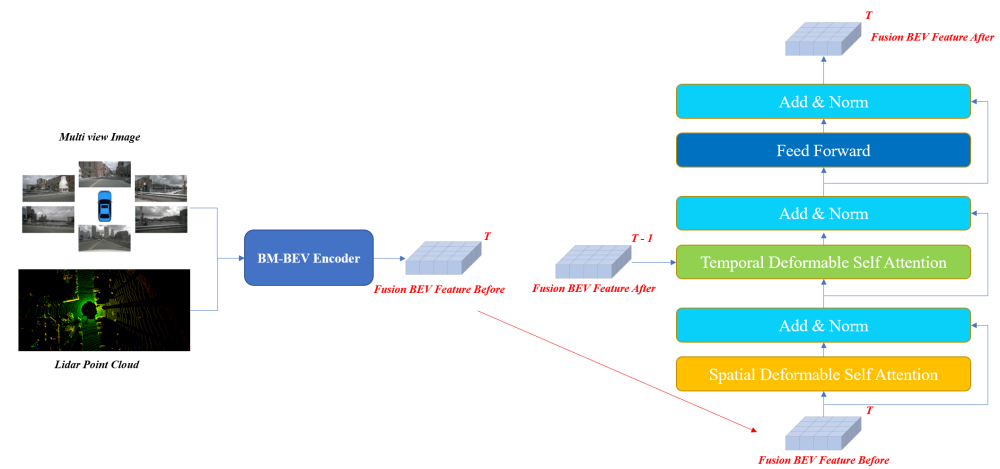
* + - **模型架构**



* **BM-BEV框架中融合时序信息**

* + **参考关联论文，针对现有BM-BEV框架的特点，对Temporal Model做了3个方案**

* + - **基于Transformer来融合时序上的信息，具体架构如下所示**

* + 

* + - * 优点

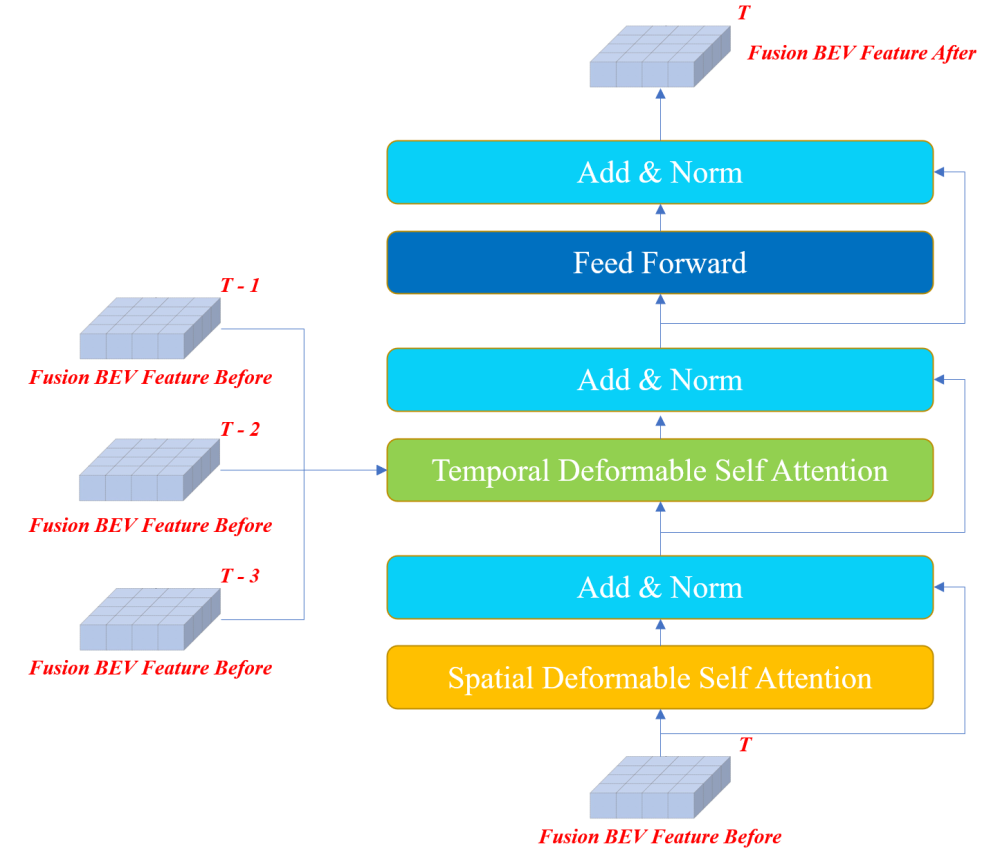
* + - * + 将当前帧增强后的BEV特征作为下一帧的History BEV特征，使用增强型BEV特征有助于时序信息的更好融合

* + - * 缺点

* + - * + 训练阶段与测试阶段存在一定的GAP，随着时间序列长度的变大，可减小GAP

* + - * + 多帧的时序信息计算很慢，生成每个时刻的帧信息无法并行，无法计算加速

* + - **改进Transformer时序融合方法，可以并行计算进行加速，架构如下图所示**

* + 

* + - * 优点

* + - * + 时序History特征生成可以并行加速计算以节省计算资源

* + - * + 训练阶段与测试阶段不存在GAP

* + - * 缺点

* + - * + 特征融合性能似乎没有第一种方法那么强

* + - **基于LSTM/RNN的方法来融合时序特征**

* + - * **将第2种方法中使用的Transformer更改为LSTM/RNN来进一步提升模型性能**

* + - * **优点**

* + - * + **速度很快**

**基于时序方案二的实验:**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Models** | **Input Size** | **Image Encoder** | **View Transformer** | **Voxel Encoder** | **Points**  **Middle Encoder** | **BEV Encoder** | **BEV Neck** | **Lr-Config** | **MAP** | **NDS** | **MMDet3D-Config** | **MMDet3D-Log** |
| LSSDet-Mixup | 256,704 | Efficientnet-b3 | Lift-Splat-64 | PillarFeatureNet | PointPillarsScatter | SECOND | SECFPN | cyclic（25） | 45.26 | 54.68 | [📎lssformerseq\_efficientnet-b3\_pointpillars\_second\_secfpn\_circlenms\_256\_704\_cyclic\_40e\_nus.py](https://yuque.antfin.com/attachments/lark/0/2022/py/21256453/1657187245781-455888e4-c5da-4847-bca9-d47a0d2fb9ec.py) | [📎20220630\_205112.log](https://yuque.antfin.com/attachments/lark/0/2022/log/21256453/1657187245798-bb5b45a3-0968-4ab6-be98-5dbeff4b8e01.log) |

在训练后期NDS比较稳定，MAP会下降，猜测由于Backbone过于弱，导致加入Transformer之后过拟合了，这里需要实验分析一下定位问题