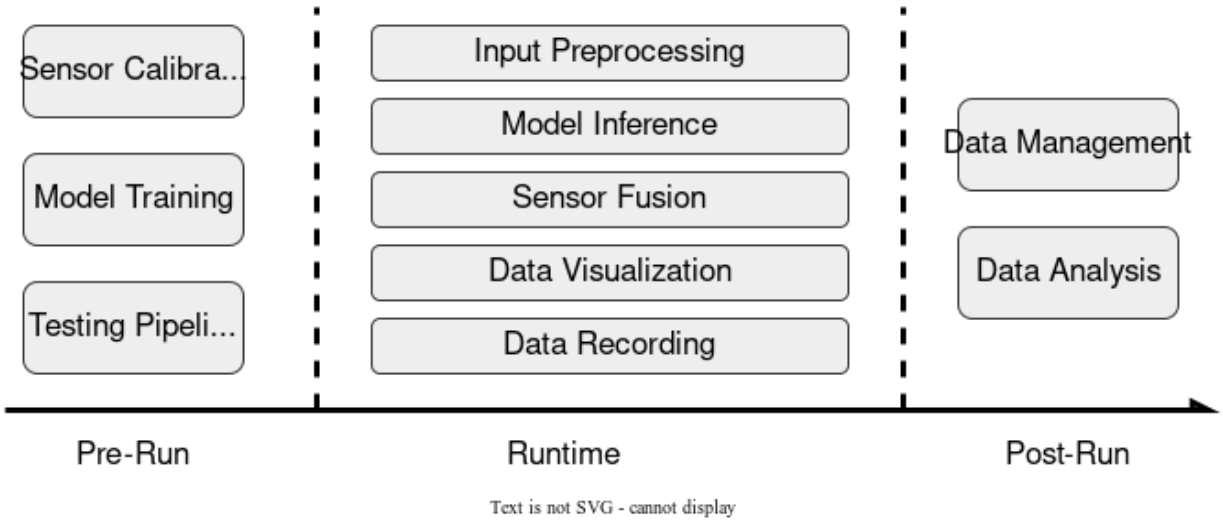
Perception Architecture of our stack



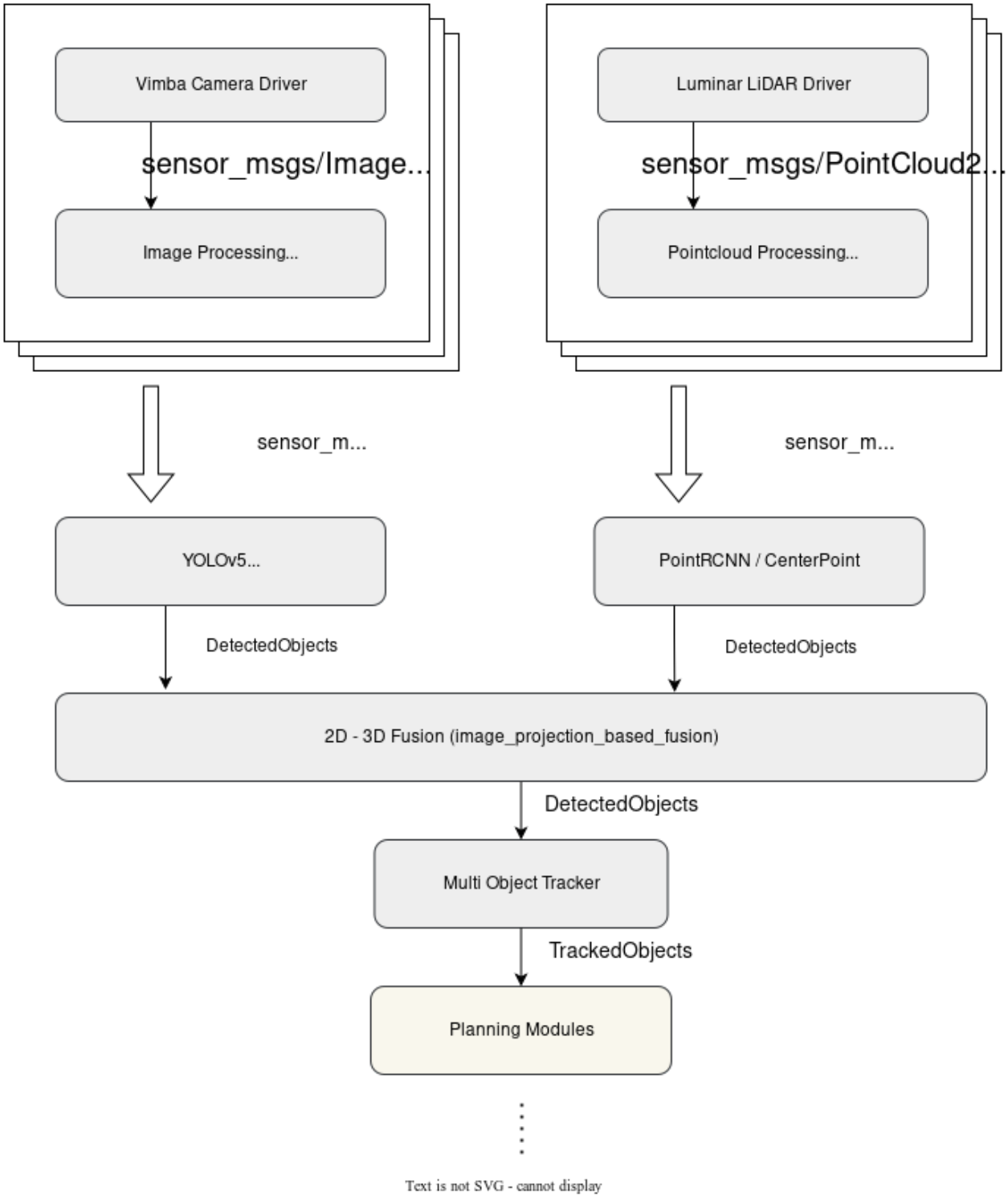
The perception stack is a heavy iterative cycle composed off of 3 stages

- Pre Run
- Runtime
- Post Run

The pre run stage is composed off sensor calibration, model training and model testing. Currently we use Autoware for all these steps. The Autoware module allows us to take raw data and calibration both the Camera intrinsics as well as the Lidar and Camera extrinsics. These are then trained through 2 models supported by Autoware. A Yolov5 Network and CenterPoint network. Once done we test these models against verification datasets. They are then tested in runtime over ROS2 bags to see their performance

Runtime is explained below. Once runtime is done we collect the data we encountered during that run and further train the models if needed. We also analyze where the models had a hard time detecting and where it went well. Finally we upl a common NAS which can be accessed by anyone in the team

latest



During runtime, the input raw data from the car is filtered, corrected and preprocessed before being sent into the models to get detections. Once we have detections they are fused using a 2D-3D associator from Autoware. They are then sent into the Autoware multi object tracker which provides stability to the detections as well as allows us to infer the speed of the vehicles based off of our own odometry

From then these tracked objects are sent into the planning module. In the future there will be a prediction module in b/w the Perception and the Planning that will also predict the car behavior so that the Planner can use that to better plan maneuvers