Baseline model – YOLOv9c (YOLOv9 Compact)

Approach: 3 - Phase Approach for improving the model

1. Fine tune the model on custom datasets like BDD10k (Berkely Deep Drive 10k) to improve model performance in low light and clutter conditions. Also addresses domain shift by incorporating in – street view of vehicles and pedestrians rather than CCTV footage.
2. Using Hybrid Feature Fusion with VGG16 to have better accuracy in crowded conditions by incorporating context – aware object detection through the use VGG16’s convolution base attached with YOLOv9’s detection head. This ensures that the final model not only gets better at detecting objects but also improves its abilities to understand the context and the environment, which helps it in recognizing the object better.
3. Improving the distance estimation algorithm by using deepSORT instead of the baseline monocular pinhole camera model ensures that final output processing of the input video is smooth. deepSORT uses ID based tracking and improves object tracking via a lightweight CNN.

*The proposed 3 – Phase Approach:*

