

## A. Effects of NMS

Here we demonstrate the effects of NMS on the sparse predictions of our Voxel-DETR and ConQueR in Table 12. Improper NMS configurations (*e.g.*, score and IoU thresholds) can cause performance degradation for all categories. And we find that the NMS configuration adopted by dense detectors (*i.e.*, score threshold 0.1, IoU threshold 0.7) performs the best. For small and densely populated categories such as pedestrians, NMS can bring noticeable performance gains, which can be observed from our baseline Voxel-DETR. However, for the well-trained large vehicles, NMS comes with a significant performance penalty, which in turn demonstrates the effectiveness of our sparse 3D object detection framework. For cyclists, NMS fluctuates in its effects on detection performance, which indicates that NMS is not necessarily required for this category. We conclude that the impact of NMS on detection performance is originates from our baseline Voxel-DETR and the inherent learning difficulty in data for extremely close query predictions, rather the Query Contrast mechanism.

Methods	Veh.	Ped.	Cyc.
<b>validation set</b>			
Voxel-DETR	68.2	64.7	70.1
Voxel-DETR <sub>nms</sub>	67.1 (-1.1)	67.1 (+2.4)	70.2 (+0.1)
ConQueR	70.5	68.1	73.3
ConQueR <sub>nms</sub>	69.2(-1.3)	70.1 (+2.0)	74.1 (+0.8)
<b>test set</b>			
ConQueR	73.3	68.7	71.9
ConQueR <sub>nms</sub>	72.6 (-0.7)	70.9 (+2.2)	71.7 (-0.2)

Table 12. Effects of NMS. APH/L2 results are reported. The numbers in brackets indicates increase (blue) or drop (red) in detection performance. <sub>nms</sub> denotes further conducting NMS on their corresponding sparse predictions.