

Appendix

In this appendix, we provide further details about our proposed method. Specifically, we describe the network architectures in detail in Section B and provide more quantitative results in Section A.

A. Additional Quantitative Evaluation

We provide per-frame model performance on real-world ScanNet+Scan2CAD and the synthetic dataset DYN-SYNTH. In Table 5, we show class-wise detection results evaluated as mean average precision at a 3D IoU of 0.5 (mAP@0.5). The per-frame completion performance is evaluated in Table 6 using a mean average precision metric with mesh IoU threshold of 0.25 (mAP@0.25).

B. Network Details

We detail the architecture of our network in Figure 6. We provide the convolution parameters as (n_in, n_out, kernel_size, stride, padding), where stride and padding default to 1 and 0, respectively. Each convolution (except the last) is followed by batch normalization and a ReLU.

	bathtub	bed	bookshelf	cabinet	chair	desk	sink	sofa	table	toilet	mAP
DYNSYNTH	49.3	38.4	12.5	6.3	44.1	46.8	27.6	32.3	38.4	63.1	35.8
ScanNet+Scan2CAD	38.7	-	12.9	4.6	41.2	-	-	26.4	29.2	-	25.6

Table 5. 3D Detection results on DYNSYNTH and ScanNet with Scan2CAD targets at mAP@0.5.

	bathtub	bed	bookshelf	cabinet	chair	desk	sink	sofa	table	toilet	mAP
DYNSYNTH	34.8	23.6	12.7	11.4	38.4	34.1	32.2	41.1	29.9	52.6	31.1
ScanNet+Scan2CAD	20.4	-	8.6	12.7	24.4	-	-	23.9	12.2	-	17.1

Table 6. Instance Completion results on DYNSYNTH and ScanNet with Scan2CAD targets at mAP@0.25.

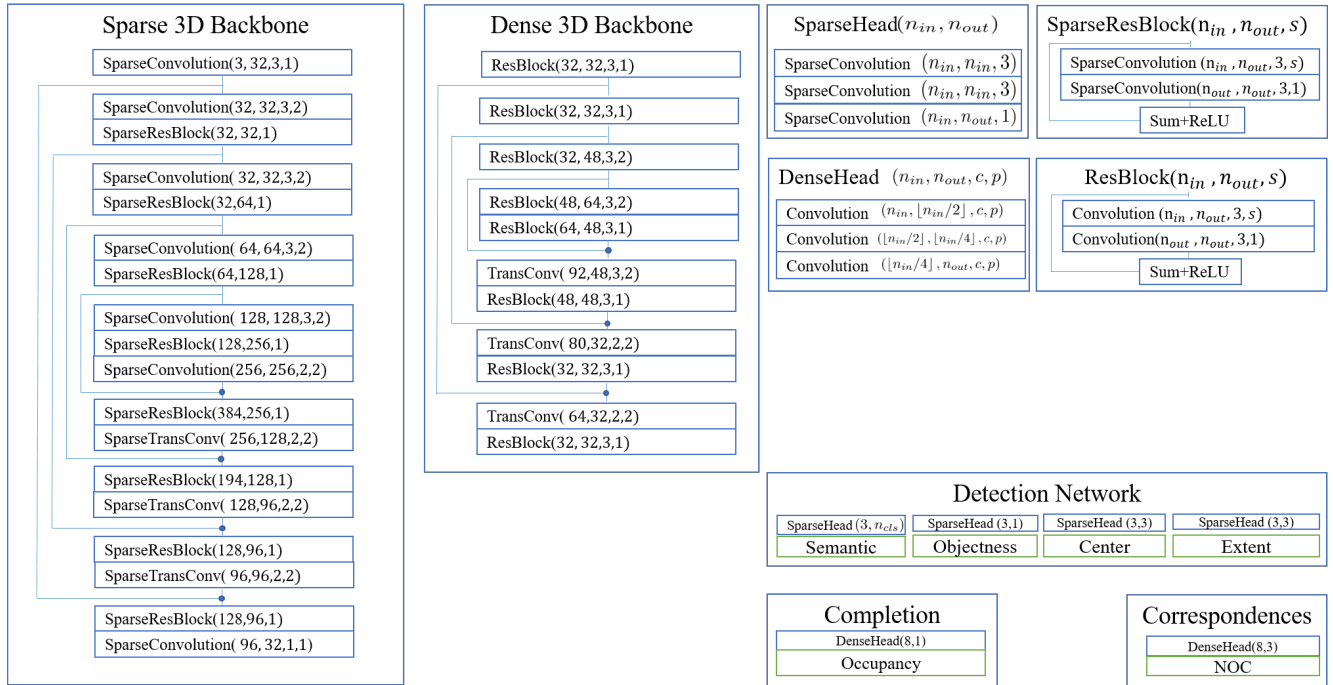


Figure 6. Network architecture specification for our approach. Dots indicate concatenation, outputs are highlighted in green.