

Supplemental Material for SG-NN: Sparse Generative Neural Networks for Self-Supervised Scene Completion of RGB-D Scans

1. SG-NN Architecture Details

Figure 1 details our Sparse Generative Neural Network specification for scan completion. Convolution parameters are given as (nf_in, nf_out, kernel_size, stride, padding), with stride and padding default to 1 and 0 respectively. Arrows indicate concatenation, and \oplus indicates addition.

Each convolution (except the last) is followed by batch normalization and a ReLU.

2. Varying Target Data Incompleteness

Here, we aim to evaluate how well our self-supervision approach performs as the completeness of the target data

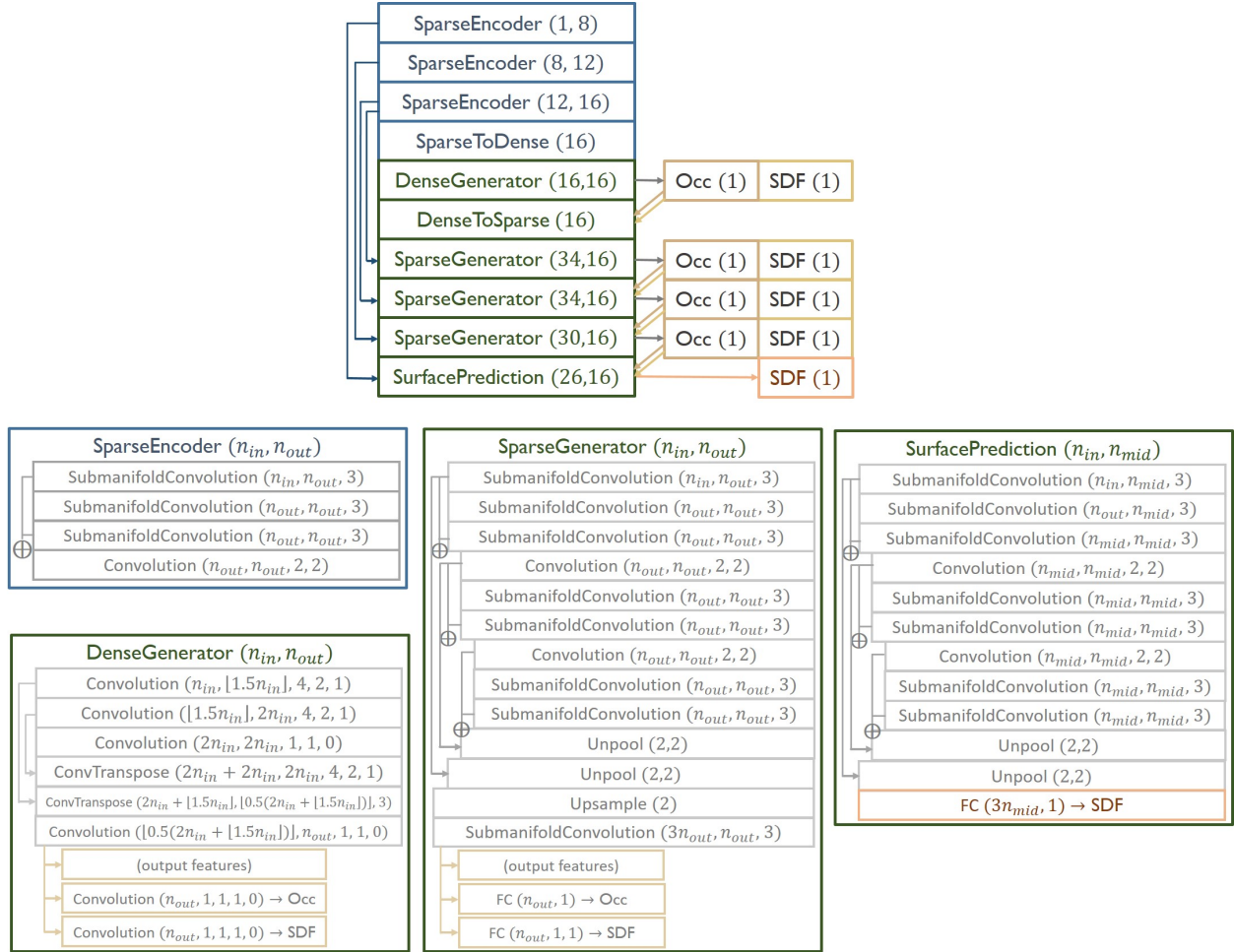


Figure 1: SG-NN architecture in detail. The final TSDF values are highlighted in orange, and intermediate outputs in yellow. Convolution parameters are given as (nf_in, nf_out, kernel_size, stride, padding), with stride and padding default to 1 and 0. Arrows denote concatenation, and \oplus denotes addition.

seen during training decreases. As long as there is enough variety in the completion patterns seen during training, our approach can learn to generate scene geometry with high levels of completeness. To evaluate this, we generate several versions of target scans from the Matterport3D [1] room scenes with varying degrees of completeness; that is, we use $\approx 50\%$, 60% , and 100% of the frames associated with each room scene to generate three different levels of completeness in the target scans, using $\approx 30\%$, 40% , and 50% for the respective input scans. We provide a quantitative evaluation in the main paper, and a qualitative evaluation in Figure 2. Even as the level of completeness in the target data used decreases, our approach maintains robustness its completion, informed by the deltas in incompleteness as to the patterns of generating complete geometry.

References

- [1] Angel X. Chang, Angela Dai, Thomas A. Funkhouser, Maciej Halber, Matthias Nießner, Manolis Savva, Shuran Song, Andy Zeng, and Yinda Zhang. Matterport3d: Learning from RGB-D data in indoor environments. In *2017 International Conference on 3D Vision, 3DV 2017, Qingdao, China, October 10-12, 2017*, pages 667–676, 2017. 2, 3

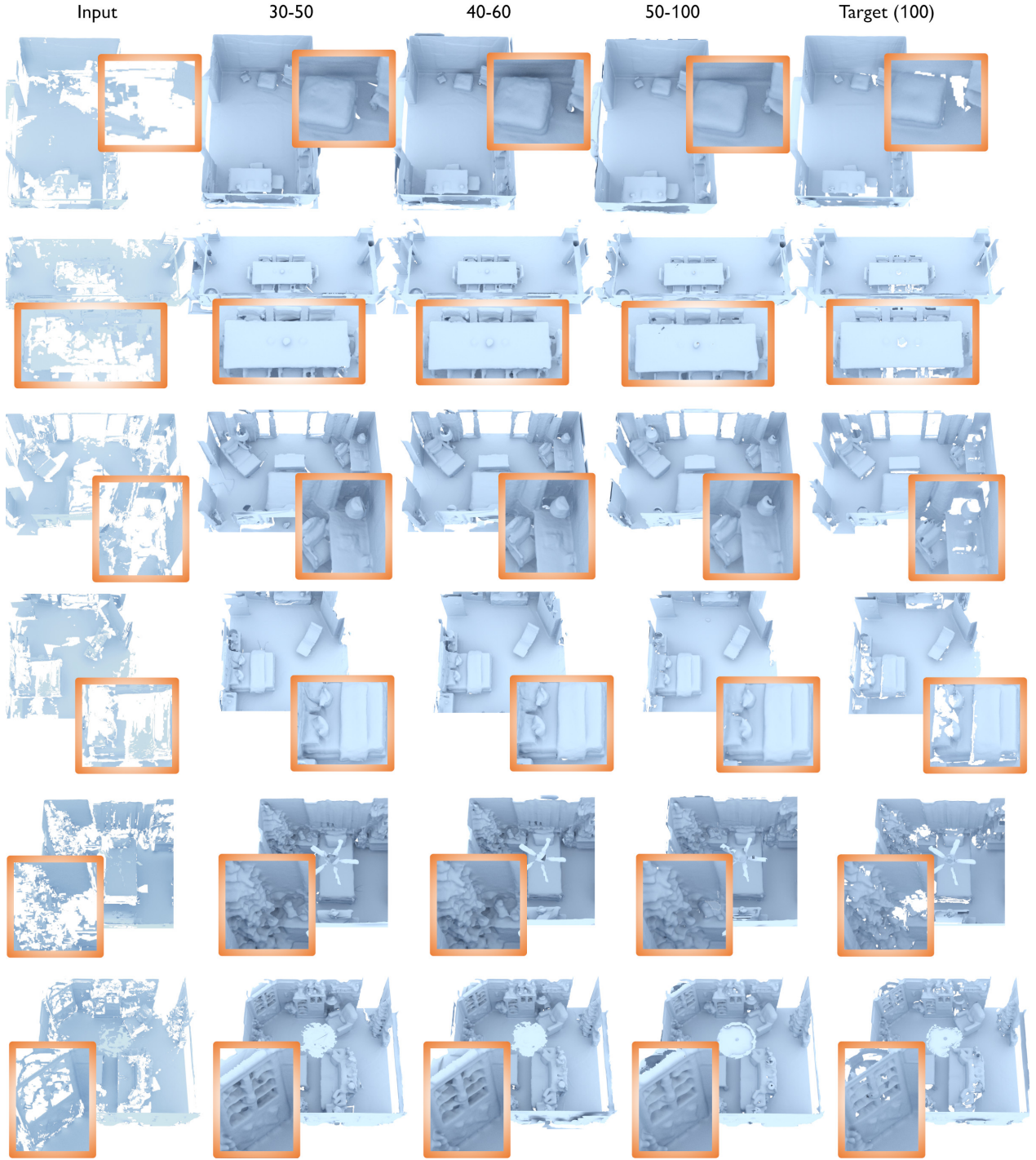


Figure 2: Qualitative evaluation of varying target data completeness available for training. We generate various incomplete versions of the Matterport3D [1] scans using $\approx 30\%$, 40% , 50% , 60% , and 100% of the frames associated with each room scene, and evaluate on the 50% incomplete scans. Even as the level of completeness of the target data used during training decreases significantly, our self-supervised approach effectively learns the geometric completion process, maintaining robustness in generating complete geometry.