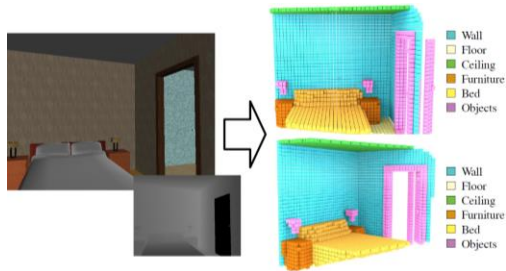


SEMANTIC SCENE COMPLETION

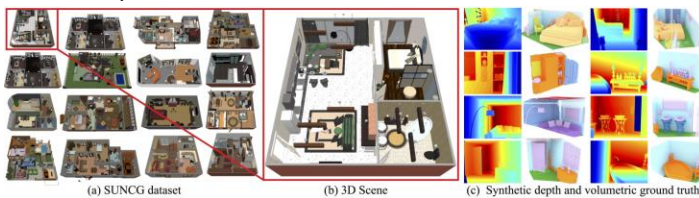


Given an RGB-D image, the goal of semantic scene completion is to infer a complete 3D occupancy grid with associated semantic labels. Previous works completely neglect the RGB channels from the input data or require a complex two step training process to merge RGB and depth data.

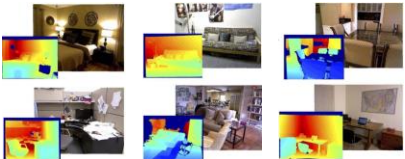
Our Edge-Net representation encodes colour information in 3D space using edge detection and flipped truncated signed distance (F-TSDF), which improves semantic completion scores especially in hard to detect classes, with an end-to-end 3D deep neural network.

DATASETS

SUNCG - Synthetic Scenes



NYUDv2 - Indoor RGB-D Scenes

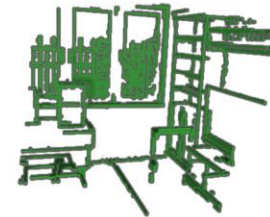


Our deep CNN is trained on SUNCG and fine-tuned on NYUDv2.

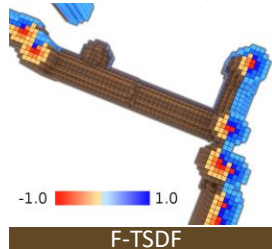
OUR APPROACH



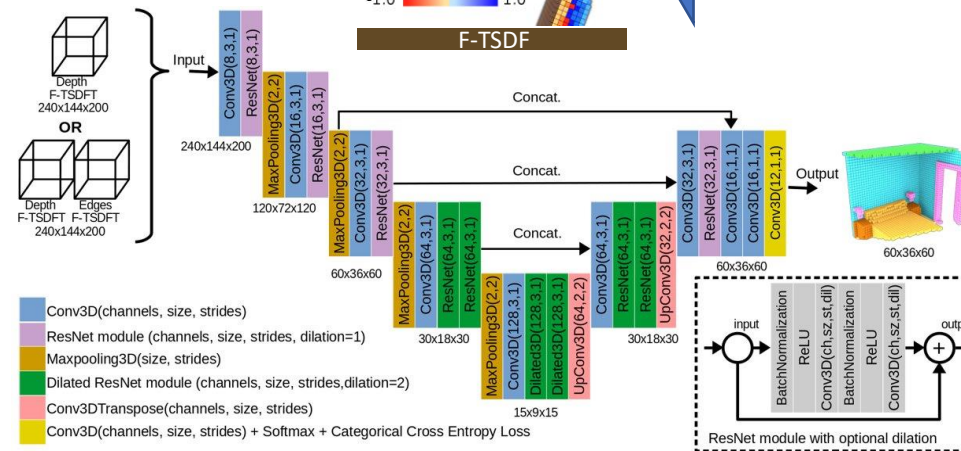
RGB edge extraction and projection to 3D



Feed CNN with 3D encoded edges and depth volumes.

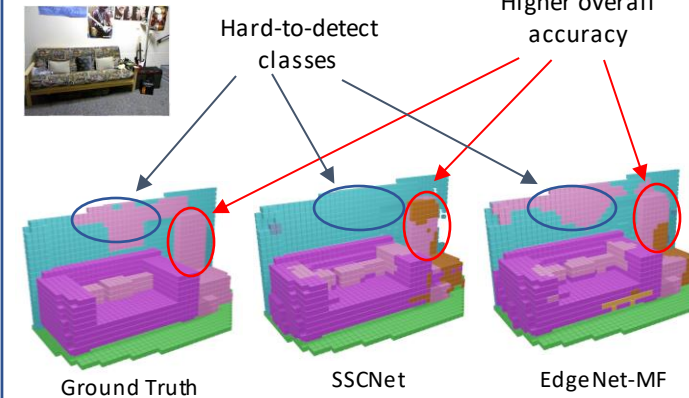


Apply F-TSDF to 3D edges



Using edges and F-TSDF, we address the data sparsity problem faced by previous solutions that tried to explore the RGB components of the RGB-D data. Our solution is an end-to-end network architecture that may be trained as a whole and achieves state-of-the-art results.

QUALITATIVE RESULTS



QUANTITATIVE RESULTS

- New state-of-the-art result on SUNCG (70.3% avg. IoU)
- Our solution surpassed previous end-to-end approaches on NYUDv2 (33.7% avg. IoU)
- EdgeNet's results are similar to non end-to-end solutions, with a much simpler training pipeline.

SUMMARY OF THE CONTRIBUTIONS

- A new end-to-end network architecture
- A new strategy to encode information obtained from RGB
- Improvement over the state-of-the-art result on SUNCG
- An efficient and lightweight training pipeline for the task