Input: Single RGBD image

Segmentation into soup of segments i.e. multiple overlapping regions
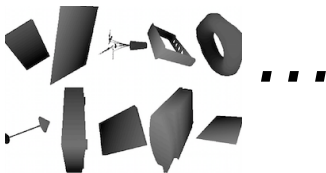
For each region...

Propose matching items from database

Align matches to region

Filter out bad matches

Database of CAD models

Combine good matches together

Output: Per-voxel occupancy prediction