Master Thesis in Visual Computing:

Semantic Understanding of 3D Scenes with Deep Learning

The goal of this thesis is to infer semantic information on 3D reconstructions obtained with a commodity RGB-D sensor (e.g., Microsoft Kinect). Here, the focus is on a data-driven technique based on deep learning that directly operates on 3D content. More specifically, the task is to develop a convolutional neural net architecture based on 3D convolutions that can be directly applied to a volumetric grid. The final goal is to predict dense voxel class labels in scanned indoor environments.

Milestones:
1) Supervised training process with ScanNet data
2) Design choices of 3D deep learning architecture
3) Results on real-world scenes

Pre-requisites: Strong C++, some scripting (Lua or Python), graphics and vision background, highly self-motivated 😊


If you are interested, please contact Matthias Niessner (matthias.niessner@tum.de).