The goal of this thesis was to investigate the neural 3D surface reconstruction from multiple views with the intent to use the resulting depth maps for bin picking. Survey of papers from 2014 to 2018 showed that none of the state of the art methods would be used to control a robot arm in our setup. Therefore we decided to create our low-level neural approach which we called the EmfNet. The network is based on a pyramidal resolution refining approach. At each pyramid's layer, there are three separate networks that take part in the computation. Each of them has a definite goal, which gives us almost complete understanding of what is going on inside the network.

The EmfNet model was partially usable, but we nevertheless extended it to EmfNet-v2. First, another measuring layer was added, which freed EmfNet from depending on an unnecessary hyperparameter. Second, we used constraints on geometry for the network not to be confused by occlusions (cases where a certain part of the surface is visible only from a single camera).

Both networks were implemented and tested on a corpus that was created as a part of this thesis. A corpus containing rendered as well as real data. The process of correspondence pairing inside the network can be observed using the visualization tool.

We designed a way how to use a robotic arm and the SMF software to quickly acquire enough data to train the model.

The best trained model is able to reconstruct 80% of the surface with less than 2 mm error under 1 second of machine time.