Autobots VIP Spring 2023

VisMan Progress Presentation 2

Anthony Miyaguchi
acmiyaguchi@gatech.edu

2023-04-04

Outline

Pick and Place Simulation

- Lights, Handy, Action
- Camera Calibration Package
- GKNet in Action

Image Detection and Segmentation

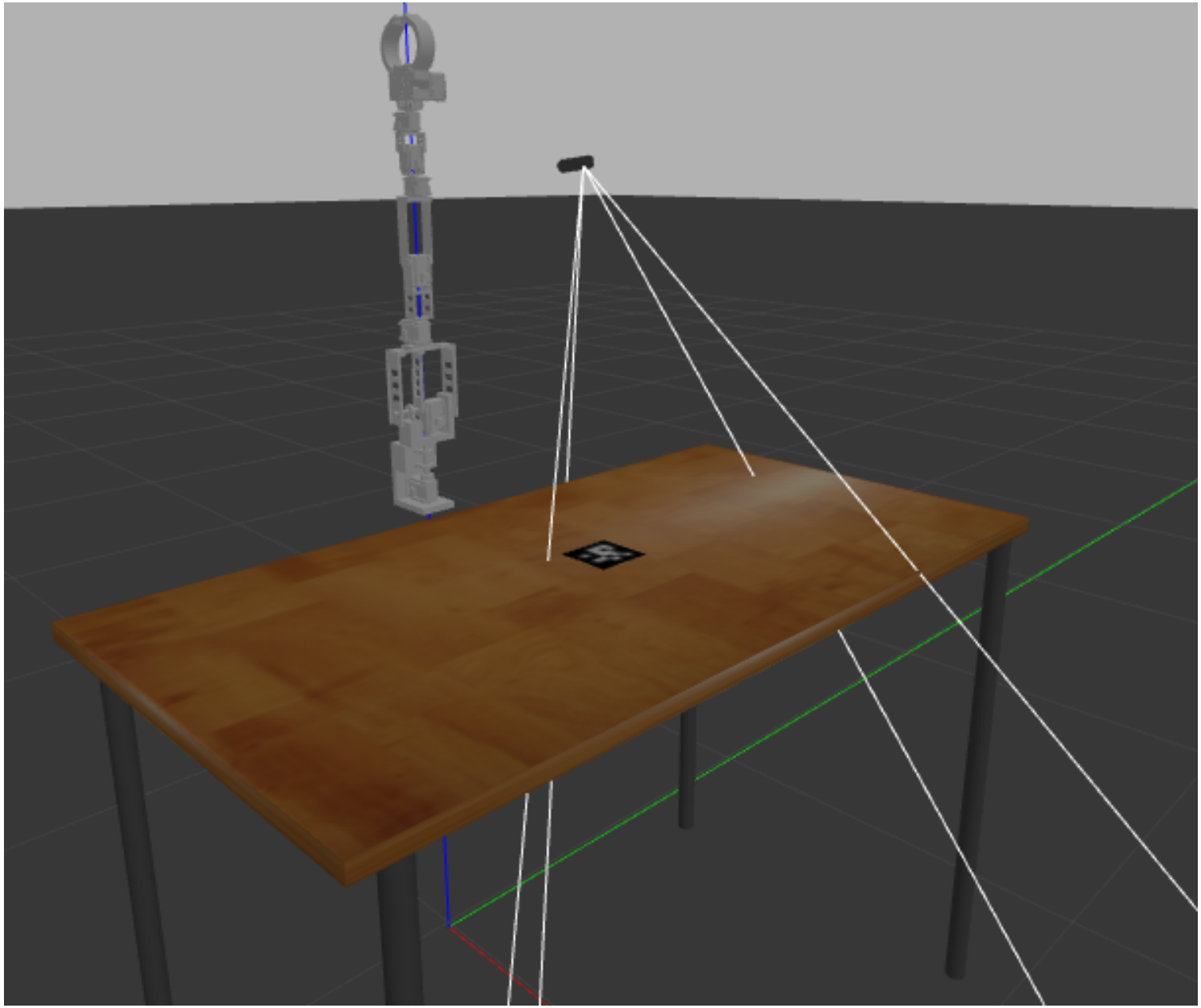
- Presented by Nick after (or before) this presentation

Lights, Handy, Action

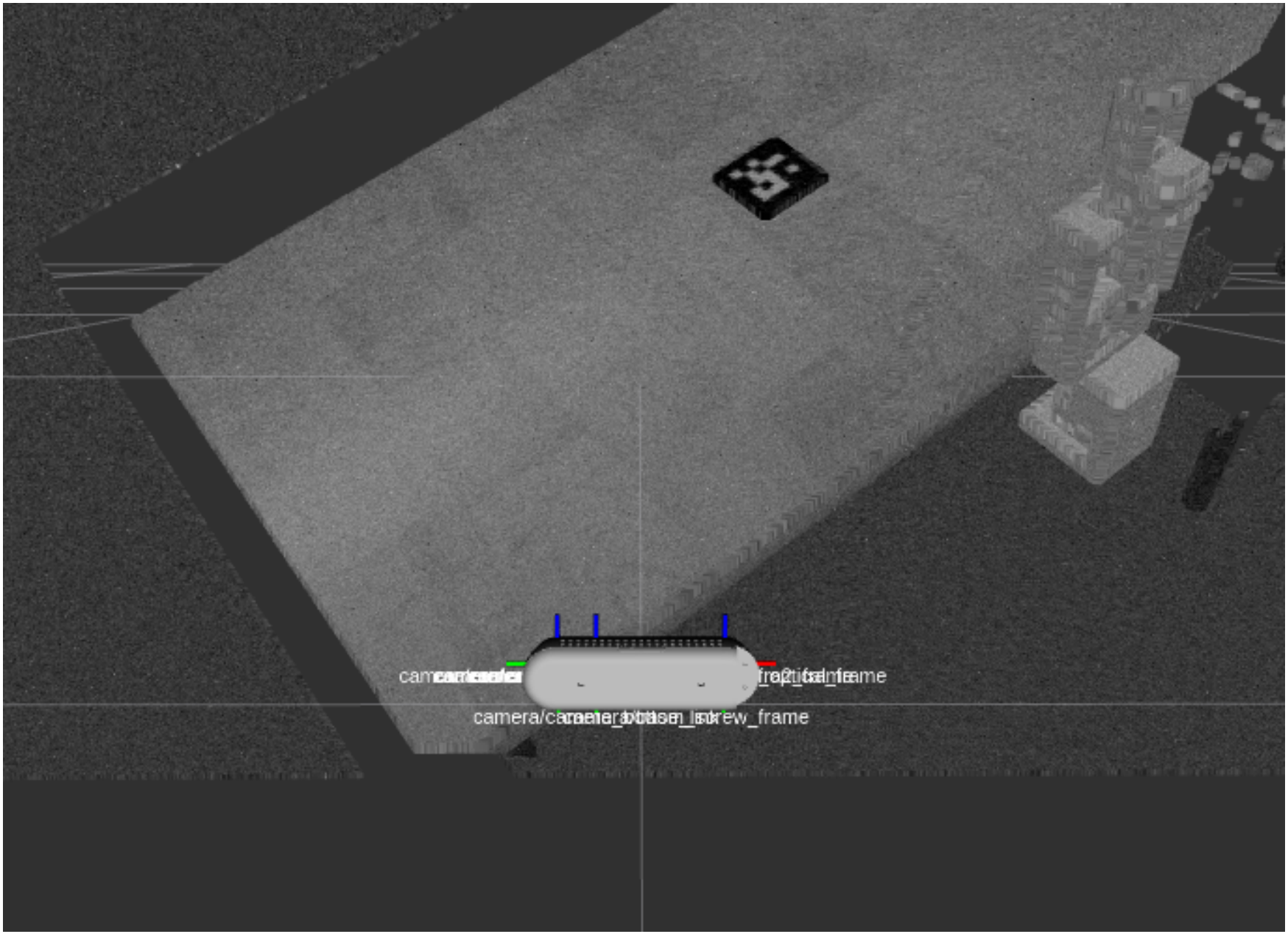
Progress

- Gazebo world with a Realsense D435, Handy, and a table.
- Gazebo is configured to work with MoveIt configured for Handy.

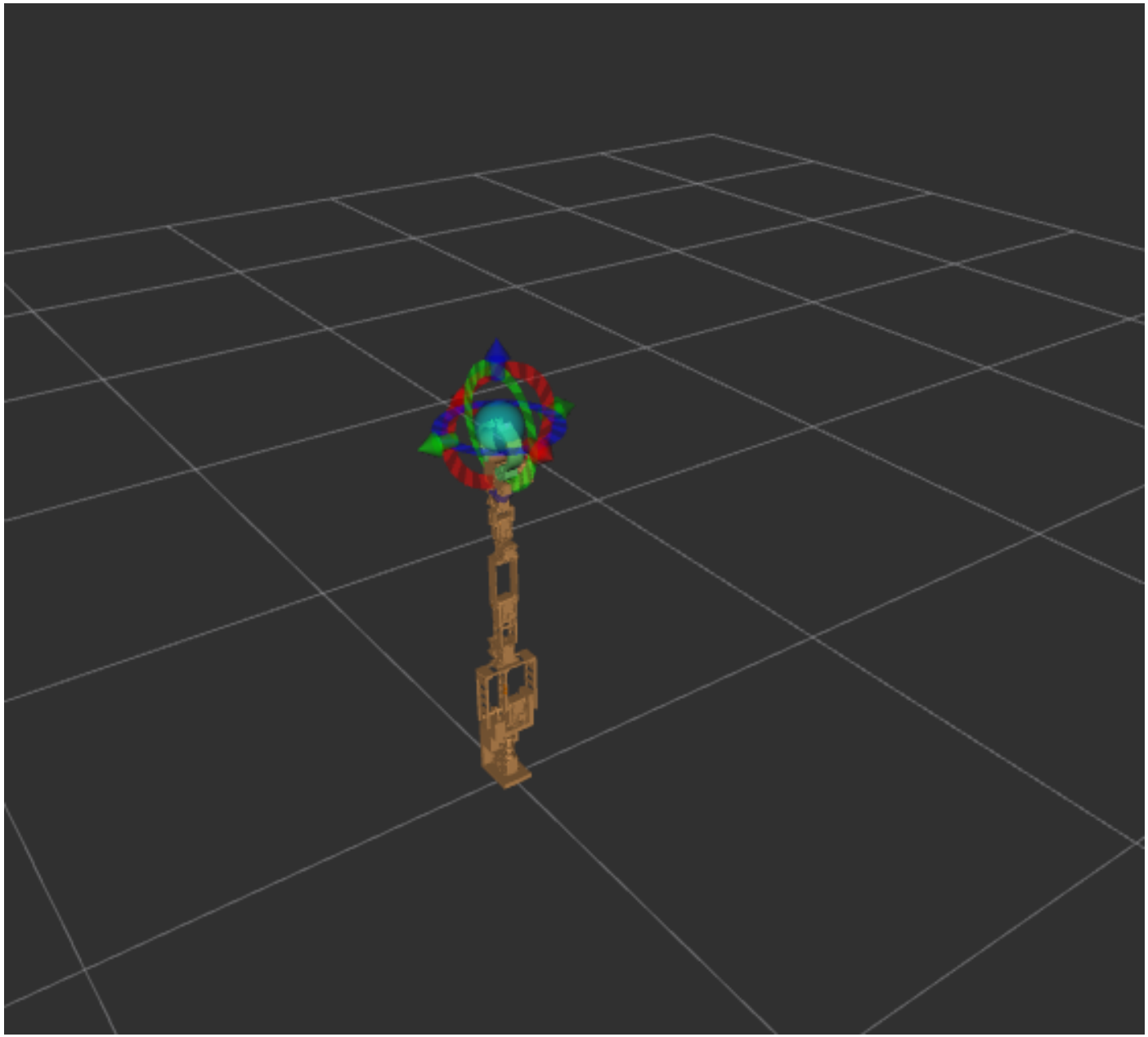
Gazebo World



Realsense D435 Point Cloud in RViz



Handy MoveIt in RViz



Demo Video

YouTube

Realsense, Handy, and Gazebo Demo Video

Camera Calibration Package

Progress

Mostly functional, reuses ivapylib/camera to determine camera extrinsics via OpenCV and a single ArUco tag. Needs more tests to ensure robustness.

Procedure

Spawn an ArUco tag with fixed world location. Then:

1. Republish the camera info topic
2. Wait until an ArUco tag is detected
3. Perform calibration; wait until transformation matrix is stable
4. Publish the transformation matrix
5. Publish calibration status

Calibration Extrinsic Publisher

Latched topic that publishes the transformation matrix from the camera to the world.

Echoing the Topic

```
$ rostopic echo /calibration/camera/camera_extrinsics
header:
  seq: 1
  stamp:
    secs: 0
    nsecs: 0
  frame_id: ''
M_CL: [
  0.9999001035866886, 0.011796350567605439, 0.007786460082420
  0.013191277641741089, -0.976675609903027 8, -0.21431458936509
  0.005076715622081074, 0.2143958934630858, -0.976733652448126
  0.0, 0.0, 0.0, 1.0
]
```

GKNet in Action

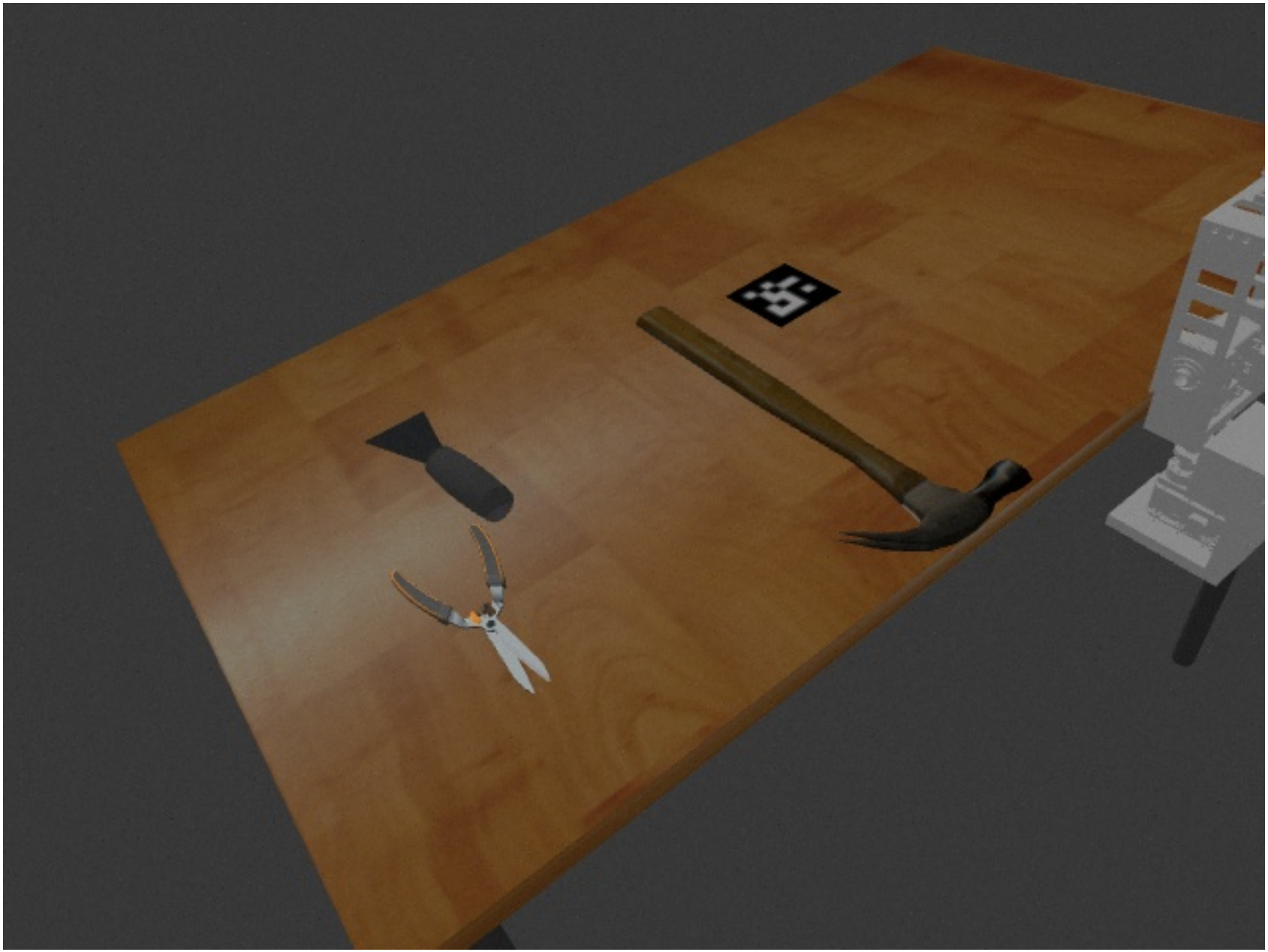
Paper Reference

Xu, Ruinian, Fu-Jen Chu, and Patricio A. Vela. "GKNet: grasp keypoint network for grasp candidates detection." The International Journal of Robotics Research 41, no. 4 (2022): 361-389.

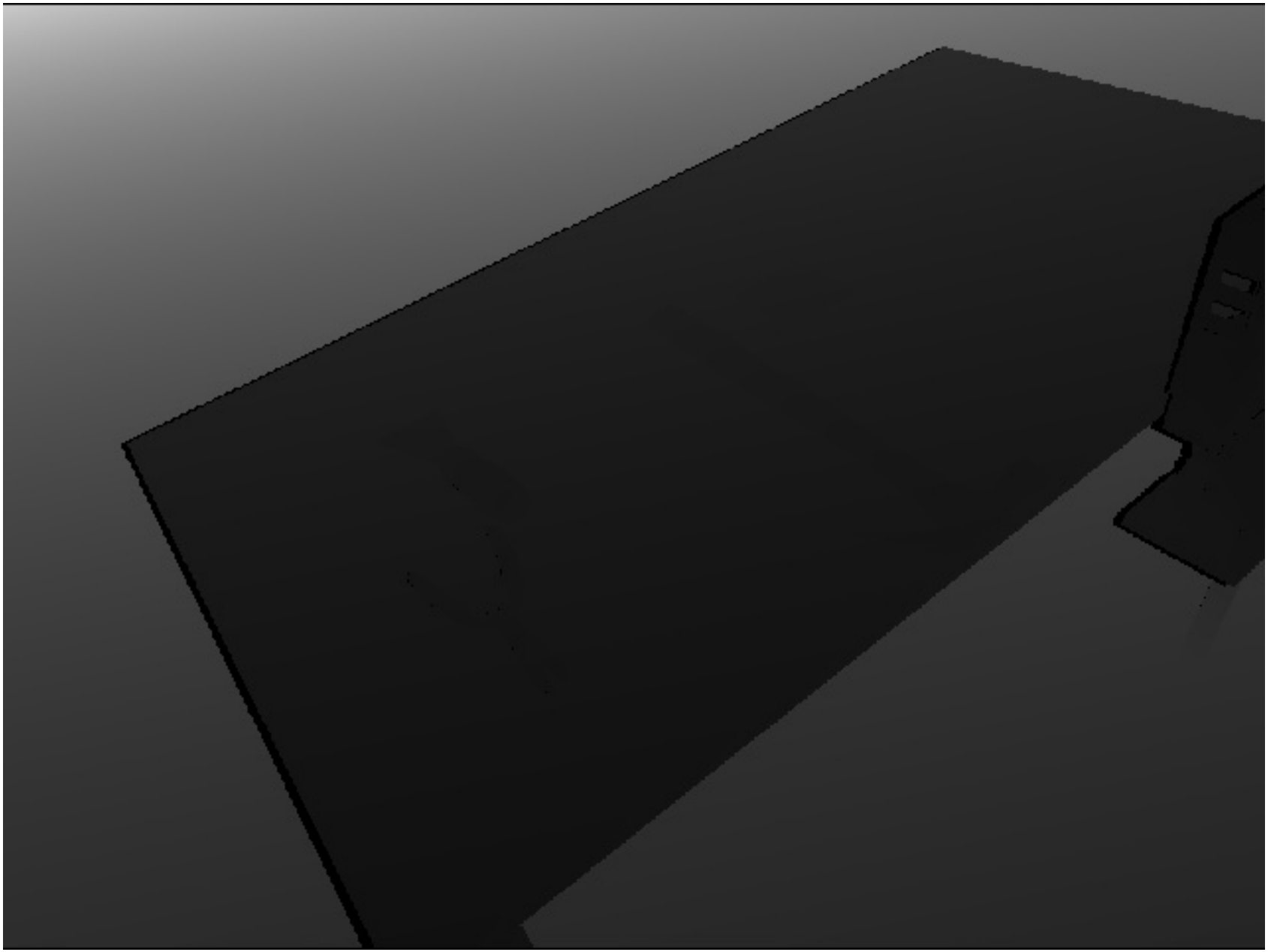
Goal

Take an image from the pick and place simulation and run it through the GKNet model.

Color Image



Depth Image



GKNet Architecture

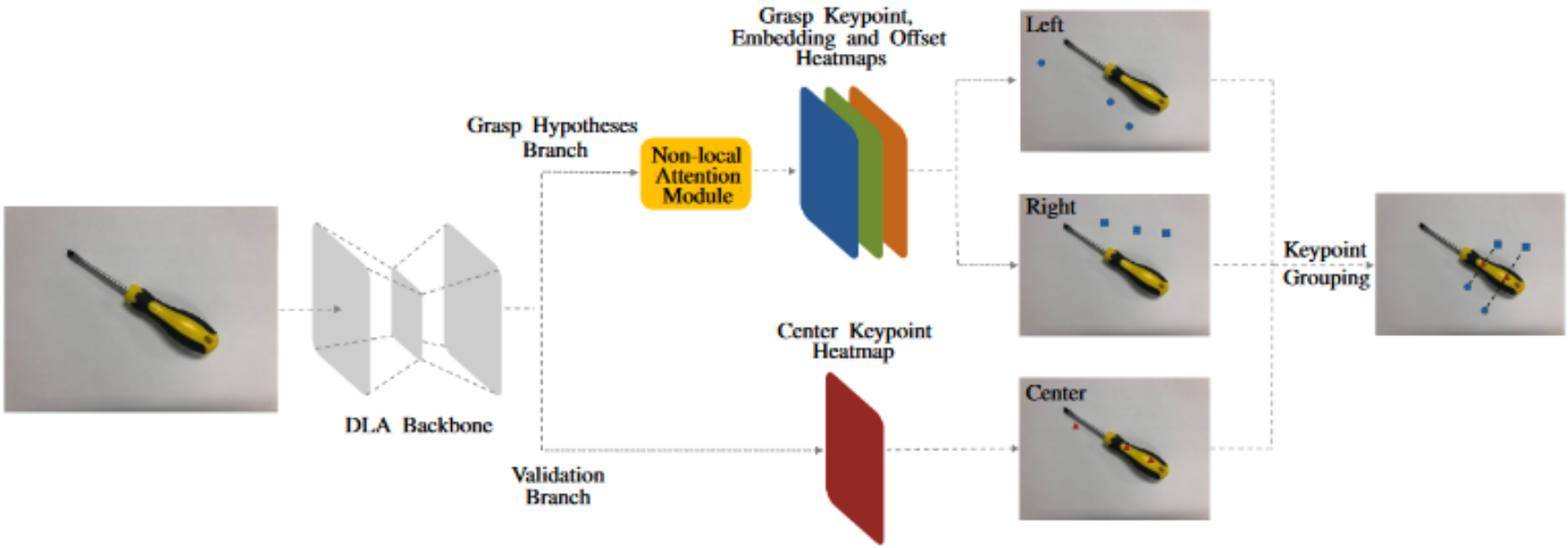
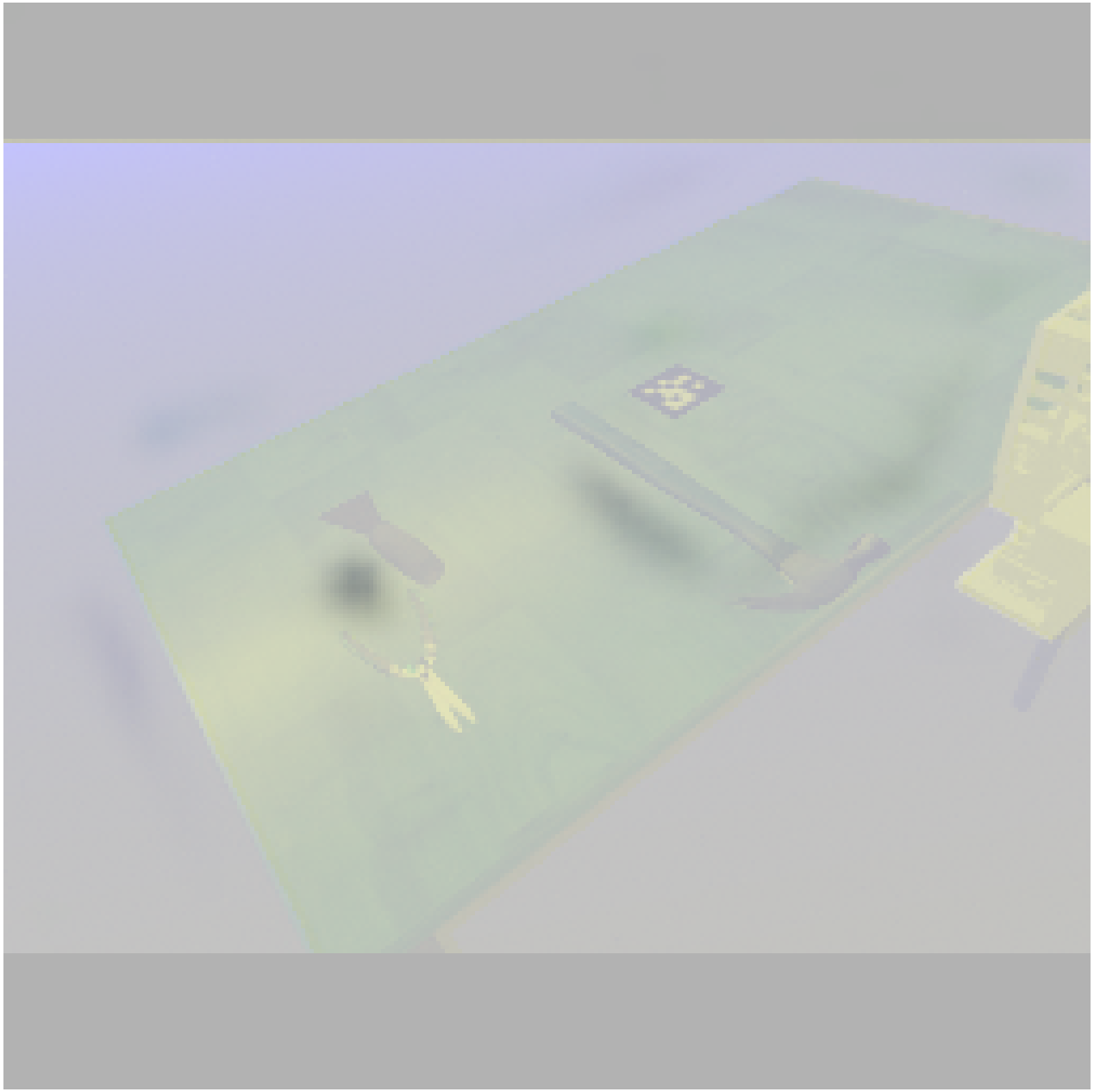


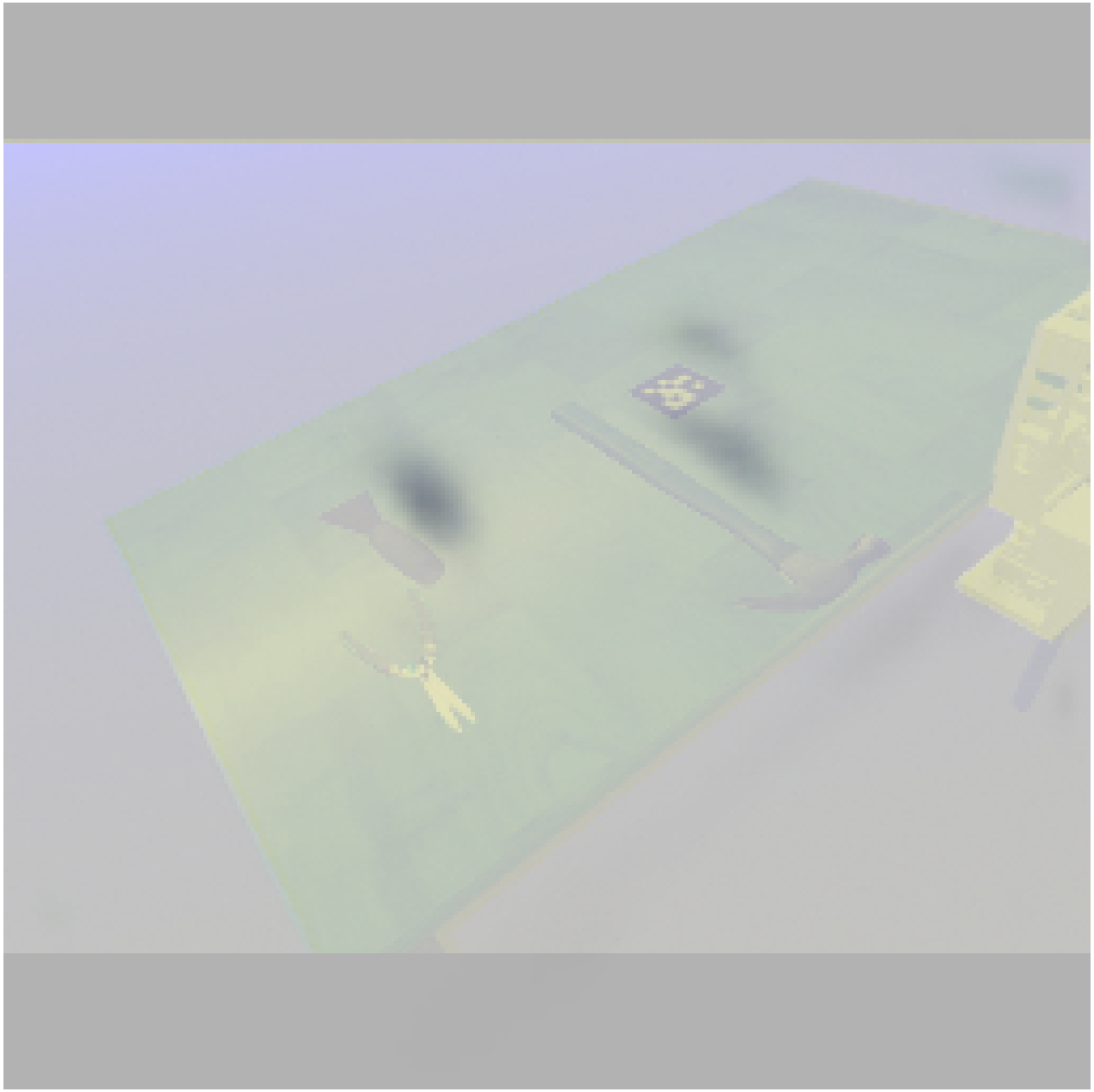
Figure 2. Architecture of GKNet. An hourglass-like backbone network is followed by two prediction branches, one for grasp keypoints and one for a center keypoint. A non-local attention module is inserted between backbone network and the prediction branch for grasp keypoints. The pair of *grasp keypoint heatmaps* with corresponding embeddings and offsets and the *center keypoint heatmap* are fed into keypoint grouping algorithm and yield the final grasp detections. Blue dots correspond to left grasp keypoints, blue rectangles correspond to right grasp keypoints and red triangles correspond to center keypoints.

Figure 6: GKNet architecture (see paper)

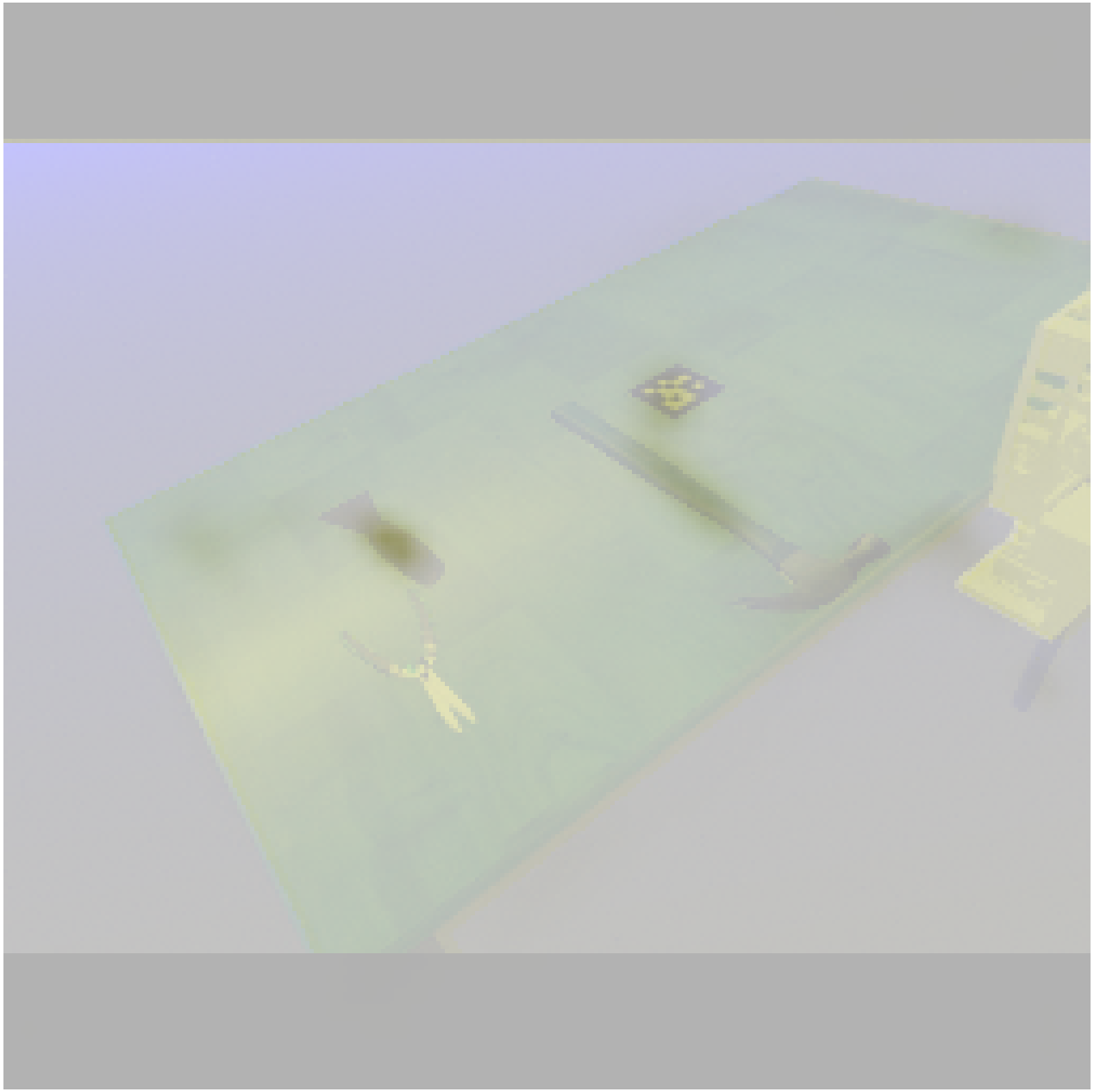
Heatmap - Left Means



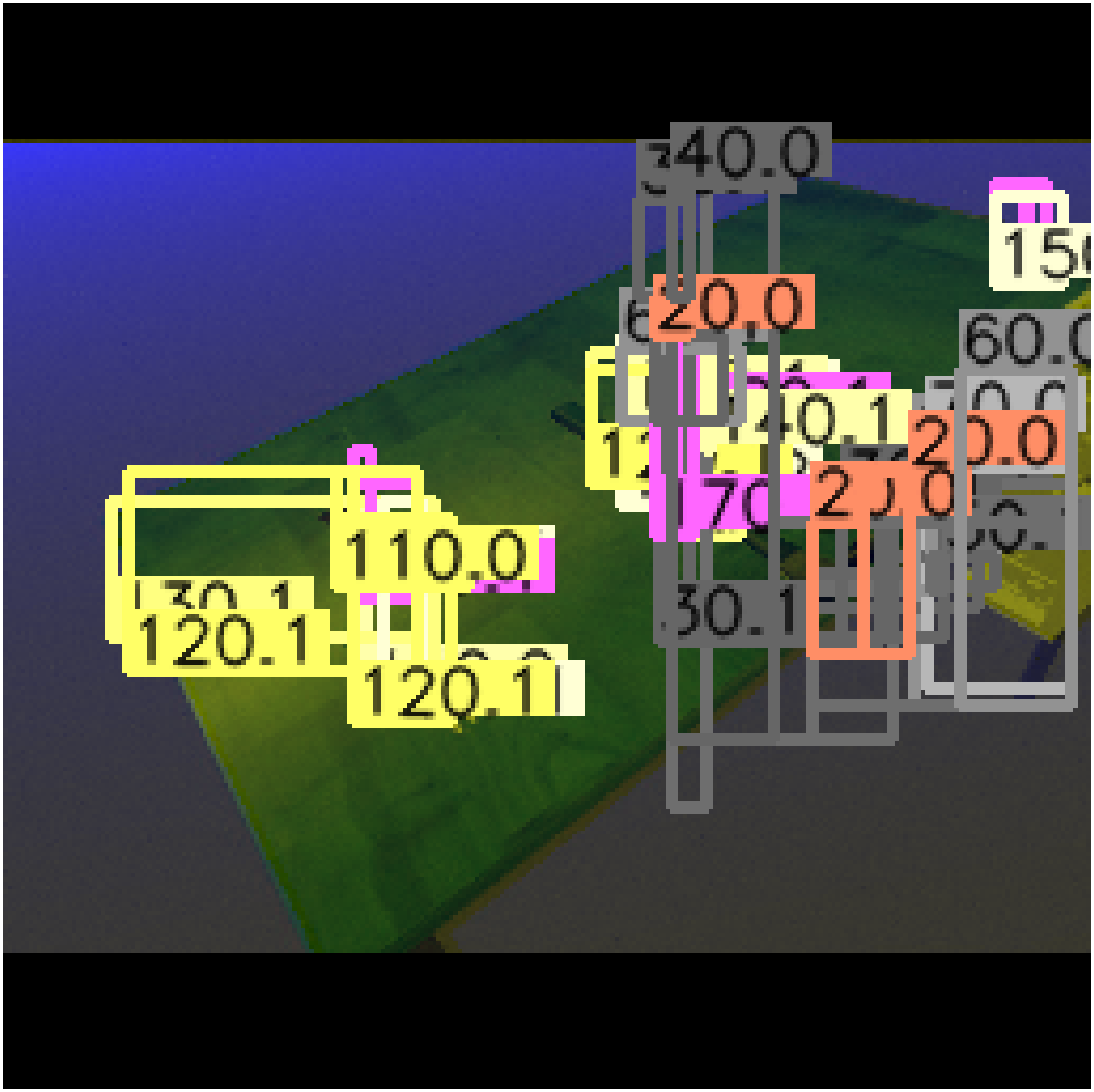
Heatmap - Right Means



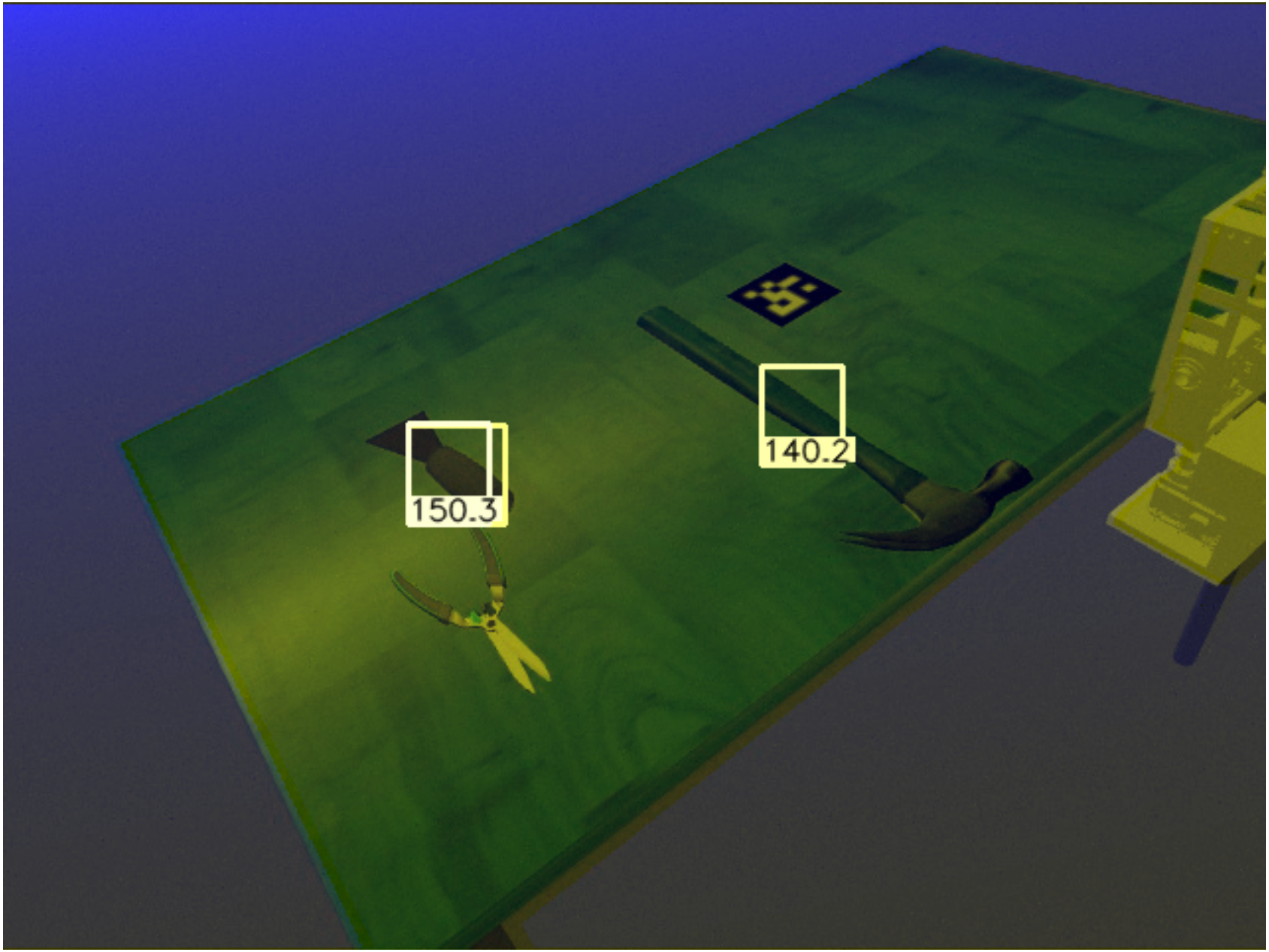
Heatmap - Center Means



Ranked Predictions



Ranked Predictions with Center Thresholding



What next?

Instead of a one-off script, the GKNet model will be integrated into a ROS node that can be used in the pick and place simulation.

Inputs

- `/camera/color/image_raw`
- `/camera/aligned_depth_to_color/image_raw`
- `/camera/color/camera_info`
- `/camera/aligned_depth_to_color/camera_info`
- `/calibration/camera/camera_extrinsics`

Outputs (TBD)

- `/gknet/predictions`
  - ranked list of keypoint pairs in camera coordinates
- `/gknet/image_annotated`
  - image with keypoints and bounding boxes drawn on it

Thank you!

Questions?