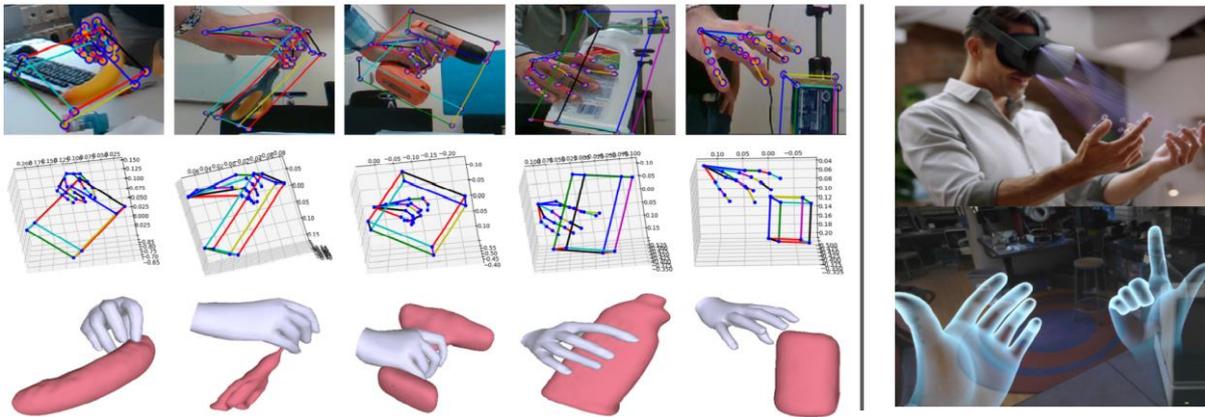


# PhD Thesis: Neural Transformers and Graph based Reconstruction of two Hands and Manipulated Object

Application



The goal of this project is to address a challenging problem of reconstructing two hands and interacting object using state-of-the-art deep learning based architectures such as transformers and graph convolutional networks. More specifically, we aim to estimate physically plausible 3D poses and meshes of two hands as well as interacting object in real-time. Recently, graph convolutional networks and transformers have proven to be effective for accurate reconstructions of human body [1][2], and pose estimation of interacting hands [1]. However, complete reconstruction (i.e., 3D poses and 3D meshes) of two hands with interacting object is a newly emerging problem which has not been well addressed in the literature. Modeling of the collisions between two hands and manipulated object is crucial for accurate reconstruction which makes this problem even more challenging. Evaluations and experiments will be performed on recently proposed datasets [1][2][3].

Our hands are the main tool that we use for interaction with objects in the real world. Thus, hand-object reconstruction from a monocular image is a very important computer vision problem. Accurately estimating hand-object meshes and poses is crucial for many practical applications including virtual and augmented reality (VR/AR), human-computer interaction, fine-grained action recognition, imitation-based learning, and telepresence. Moreover, understanding hand-object interactions are essential for developing robots that perceive and act in the world. Recently, a plethora of studies focused on hand pose estimation.

## References:

- 1- Shreyas Hampali, Sayan Deb Sarkar, Mahdi Rad, Vincent Lepetit, "HandsFormer: Keypoint Transformer for Monocular 3D Pose Estimation of Hands and Object in Interaction", In arXiv, 2020.
- 2- Taein Kwon, Bugra Tekin, Jan Stuhmer, Federica Bogo, Marc Pollefeys, "H2O: Two Hands Manipulating Objects for First Person Interaction Recognition", In arXiv, 2020.
- 3- Samarth Brahmabhatt, Chengcheng Tang, Chris Twigg, Charles C. Kemp, and James Hays, "ContactPose: A Dataset of Grasps with Object Contact and Hand Pose", In ECCV, 2020

For any questions or further questions, please contact us:

**Ahmed Elhayek**  
[Ahmed.elhayek@dfki.de](mailto:Ahmed.elhayek@dfki.de)

or

**Jameel Malik**  
[Jameel.malik@dfki.de](mailto:Jameel.malik@dfki.de)

**Released: 20.07.2021**