

We present a unified framework for understanding 3D hand and object interactions in raw image sequences from egocentric RGB cameras. Given a single RGB