

1. Predict Pixel-Aligned 3D Heatmaps

2. Predict Auxiliary EEF States

3. I.K to Robot State

