

To help bridge the gap between internet vision-style problems and the goal of vision for embodied perception we instantiate a large-scale navigation task -- Em