Training Environment (simulation) **Multimodal Sensory Input RGB** Masked mask 1 : use all modalities **Imitation Learning** Depth mask 2: ignore RGB data Proprioception Testing Environment (real robot) Masked Modality Features **RGB** mask 1 : use all modalities Policy Network Depth mask 2: ignore RGB data Actions Proprioception