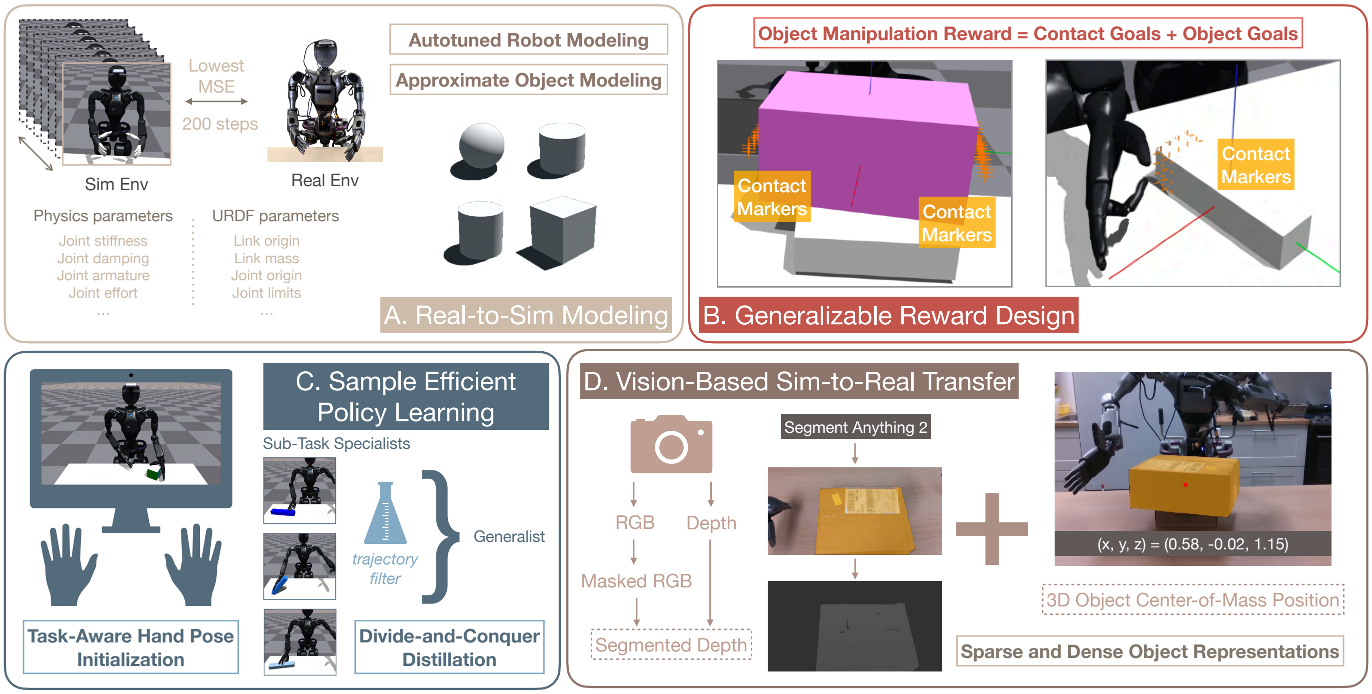

Fig. 2: **A sim-to-real RL recipe for vision-based dexterous manipulation.** We close the environment modeling gap between simulation and the real world through an automated real-to-sim tuning module, design generalizable task rewards by disentangling each manipulation task into contact states and object states, improve sample efficiency of dexterous manipulation policy training by using task-aware hand poses and divide-and-conquer distillation, and transfer vision-based policies to the real world with a mixture of sparse and dense object representations.

they do not fundamentally resolve the exploration challenge. Additionally, applying RL to solve real-world robotics also reveals important challenges that standard benchmarks in RL [5, 54] fail to capture: (1) the lack of fully or accurately modeled environments; (2) the lack of well-defined reward functions for tasks of interest.

Past works in the intersection of robotics and RL have proposed various practical techniques to alleviate these problems, such as learning from human motion data or teleoperated demonstrations [9, 45, 60, 65], real-to-sim techniques to model object and visual environments [2, 16, 17, 31, 55] and more principled ways to design rewards [37, 62]. While some of them overfit to specific tasks and settings, they point to promising directions upon which this work builds.

B. Vision-Based Dexterous Manipulation on Humanoids

Imitation learning and classical approaches. Innovations in teleoperation [10, 32, 58, 63] and learning from demonstrations [11, 28] have brought about many recent advances in vision-based dexterous manipulation [10, 28, 32, 64]. However, in practice, onboarding teleoperators to collect high-quality dexterous manipulation data remains costly, and the performance scaling with data purely collected from real-world teleoperation [27, 29, 64] suggests that the cost to reach human-level performance could be prohibitively large.

Reinforcement learning approaches. A number of existing works have successfully applied RL to solve dexterous manipulation problems with multi-fingered hands, but either assume a single-hand setup [2, 8, 17, 35, 42, 49, 57] or do not use pixel inputs as object representation [9, 20, 31]. Moreover, most of the existing works focus on a single manipulation skill, including in-hand reorientation [2, 17, 42, 57], grasping [35, 49], twisting [31], and dynamic handover [20]. The closest to our work is Chen et al. [9], but their method relies on human hand motion capture data to learn a wrist controller rather than learning the full hand-arm joint control from scratch. In addition, existing works often focus on hardware whose models in physics simulation have been more extensively tested. Our work is the first to show successful sim-to-real RL transfer of vision-based dexterous manipulation policies to a novel humanoid hardware with multi-fingered hands.

III. Challenges and Approaches

In Section I, we identify four areas of challenges in applying sim-to-real RL to dexterous manipulation and briefly describe our strategies to tackle the challenge in each area. Below, we describe our specific approaches in detail. Figure 2 shows an overview of the challenges and approaches.

A. Real-to-Sim Modeling

Simulators offer unlimited trial-and-error chances to perform the exploration necessary for RL. However, whether policies learned in simulation can be reliably transferred to the real world heavily depends on the faithfulness of modeling — both the robot itself and the environment. When applying