

# RL-GSBridge: 3D Gaussian Splatting Based Real2Sim2Real Method for Robotic Manipulation Learning

Yuxuan Wu\*, Lei Pan\*, Wenhua Wu, Guangming Wang, Yanzi Miao, Fan Xu# and Hesheng Wang#

**Abstract**— Sim-to-Real refers to the process of transferring policies learned in simulation to the real world, which is crucial for achieving practical robotics applications. However, recent Sim2real methods either rely on a large amount of augmented data or large learning models, which is inefficient for specific tasks. In recent years, with the emergence of radiance field reconstruction methods, especially 3D Gaussian splatting, it has become possible to construct realistic real-world scenes. To this end, we propose RL-GSBridge, a novel real-to-sim-to-real framework which incorporates 3D Gaussian Splatting into the conventional RL simulation pipeline, enabling zero-shot sim-to-real transfer for vision-based deep reinforcement learning. We introduce a mesh-based 3D GS method with soft binding constraints, enhancing the rendering quality of mesh models. Then utilizing a GS editing approach to synchronize the rendering with the physics simulator, RL-GSBridge could reflect the visual interactions of the physical robot accurately. Through a series of sim-to-real experiments, including grasping and pick-and-place tasks, we demonstrate that RL-GSBridge maintains a satisfactory success rate in real-world task completion during sim-to-real transfer. Furthermore, a series of rendering metrics and visualization results indicate that our proposed mesh-based 3D GS reduces artifacts in unstructured objects, demonstrating more realistic rendering performance.

## I. INTRODUCTION

Learning robotic action policies in simulation and transferring them to real-world represents an ideal robotic learning strategy that balances both the cost and safety of the learning process. However, a significant bottleneck is the reliability of sim-to-real transfer, which impacts the potential of the entire framework towards substantial challenges.

With the continuous development of the simulation-to-reality(Sim2Real) field, extensive work is advancing progress from multiple perspectives [1], [2], [3], [4]. However, most Sim2Real methods attempt to expand the distribution of training data to cover various situations that may arise in

This work was supported in part by the Shenzhen Science and Technology Program under Grant KJZD20230923114812027. (Corresponding Author: Fan Xu and Hesheng Wang)

\* Equal contributions

# Co-corresponding authors

Y. Wu, W. Wu, F. Xu, and H. Wang are with the Shenzhen Research Institute of Shanghai Jiao Tong University, Shenzhen 518000, China.

Y. Wu, F. Xu, and H. Wang are also with the Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China. (e-mail: furygreen@sjtu.edu.cn; xufanlyra@sjtu.edu.cn; wanghesheng@sjtu.edu.cn) W. Wu is also with MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, Shanghai 200240, China.

L. Pan and Y. Miao are with the School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221100, China.

G. Wang is with the Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, U.K. (e-mail: gw462@cam.ac.uk)

Code is available at <https://github.com/IRMV-Manipulation-Group/RL-GSBridge>

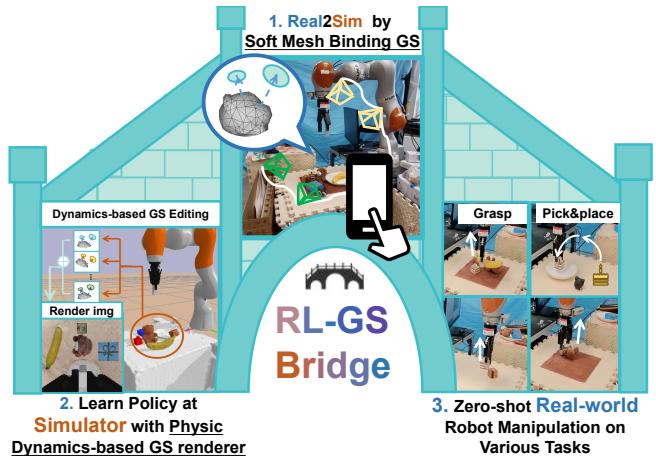


Fig. 1. Pipeline of RL-GSBridge. (1) Real2Sim Environment Transfer. Real-world scenarios is reconstructed through a novel soft mesh binding GS model. (2) Learn Policy at Simulator with GS Render. With physical dynamics-based GS editing, RL policies learn through realistic rendered images in simulation. (3) Zero-shot Real-world Robot Manipulation. We directly apply the policy to real-world tasks without fine-tuning.

reality, or to train a highly generalized large model to learn knowledge for different tasks. This significantly increases the difficulty in training stage.

To avoid additional training burden and achieve ideal Sim2Real performance on specific tasks, a novel framework is needed. Recently, advances in radiance field-based reconstruction methods [5], [6], [7], [8] provide new directions for Sim2Real training. Based on a simple idea—using radiance field reconstruction to create a visually realistic robot training environment—can we achieve satisfactory Sim2Real performance? For this purpose, we design a Real2Sim2Real visual reinforcement learning framework, RL-GSBridge, that bridges the real-to-sim gap by 3D Gaussian splatting, as shown in Fig. 1. Utilizing 3D Gaussian splatting and editing techniques, RL-GSBridge provide a ‘virtual-reality’ simulation platform for policy learning.

For vision-based robot tasks, it is crucial to avoid illusions caused by inconsistencies between visual perception and contact geometry. Thus, it is required to ensure an accurate geometric representation of the model while achieving more realistic rendering results. GaMeS [9] has drawn our attention as it is a mesh-based Gaussian splatting (GS) method, which ensures that the optimization of GS units is performed within the geometric mesh model. However, it enforces Gaussians to be aligned with the mesh grid planes, which could be considered as ‘hard mesh binding’, thereby limiting the

flexibility of GS units. To address this, we propose a soft mesh binding method for Gaussian Splatting, which could further enhance rendering quality while preserving editing capabilities for both objects and the background.

Physics-based dynamics simulation is also a crucial and challenging aspect of sim2real. To tackle this, an off-the-shelf physics simulator is used to provide dynamic changing information. The GS editing process simultaneously updates the scene, ensuring that the rendering results align with physical interaction processes.

Based on the designed simulation platform, we train robotic manipulation policies by deep reinforcement learning methods. The model is trained on grasping and pick-and-place tasks across scenarios that include diverse textures, geometric shapes, and patterned desktop backgrounds.

Under the RL-GSBridge framework, the policy shows a minor variation in success rates during Sim2Real transfer. This means that the policy maintains effective performance on real-world tasks, reflecting a strong ability to generalize from simulation to real-world environments.

To summarize our contributions:

- **A Novel Sim2Real RL Framework:** Leveraging the high-fidelity rendering of 3D GS and the convenience of modeling scenes with only consumer-grade cameras.
- **A Soft Mesh Binding GS Modeling Method:** Proposing a soft mesh binding strategy to replace the hard mesh binding baseline, enhancing the flexibility and render quality.
- **Physical Dynamics-Based GS Editing:** Integrating dynamics signals from the simulator to edit 3D GS models, reflecting realistic physical robotic interactions.
- **Validation on Real Physical Robots:** Testing the RL-GSBridge framework on physical robots through grasping and pick-and-place tasks in real-world scenarios with complex textures and geometries.

## II. RELATED WORK

### A. Sim2Real Transfer in RL

RL constructs an interactive learning model in which an agent learns to maximize rewards through trial and error. By incorporating deep learning, deep reinforcement learning(DRL) enhances the framework's ability to tackle more complex tasks [10]. DRL has shown impressive performance across various domains, including games [11], finance [12], autonomous driving [13], and robotics [14]. However, most applications are restricted to virtual environments due to real-world constraints related to safety, efficiency, and cost.

To improve the feasibility of deploying models in the real world, many researchers strive to bridge the gap between simulation and reality[15]. These methods include domain randomization [16] and domain adaptation [17], [18]. Domain randomization involves varying task-related parameters in the simulation to cover a broad range of real-world conditions, while domain adaptation focuses on extracting a unified feature space from both sources. Higher-level learning methods include metalearning [19] and distillation

learning [20]. Meta-learning aims to teach robots the ability to learn new tasks, whereas distillation learning trains a student network using knowledge from expert networks.

Our approach aligns with the first category of gap-bridging methods, but with a unique twist. We use soft mesh binding GS to create realistic simulation environments for robot training. Typically, achieving such fidelity requires expensive 3D scanning equipment or CAD expertise. In contrast, GS models visually realistic simulation environments using only multi-view images captured with consumer-grade devices.

### B. Radiance Field in Robotics

Neural Radiance Fields (NeRF) [5] is an implicit representation technique for novel view synthesis. It optimizes the parameters through multi-view images, and allows for the synthesis of any target views using volumetric rendering. NeRF's representation has been applied to various tasks, including Simultaneous Localization and Mapping (SLAM) [7], [21], [22], scene reconstruction [23], scene segmentation [24], navigation [25], and manipulation [26].

NeRF-RL [27] treats the novel view synthesis task of NeRF as a proxy task, where the learned encoding network is directly used as feature input for reinforcement learning. Y. Li *et al.* [28] uses NeRF as a perceptual decoder for the hidden states of a world model, training an additional dynamics estimation network to predict future state changes. NeRF2Real [29] learns real-world contexts by converting background meshes into a simulator, and trains robots to perform visual navigation, obstacle avoidance, and ball-handling tasks. However, the training and rendering speed of Vanilla NeRF has consistently been a bottleneck limiting its further deployment in practical applications.

3D GS [30] is an explicit radiance field method that directly updates the attributes of each 3D Gaussian component to optimize scene representation. 3D GS employs a splatting technique, for rendering, and achieve extremely fast training efficiency through CUDA parallel technology. Moreover, compared to the implicit representation of NeRF, the explicit representation facilitates tracking dynamic scene modeling and editing scene content.

Many robotics works also incorporate 3D Gaussian techniques for perception and motion learning [31]. ManiGaussian [32] uses 3D GS as visual and dynamic scene representation for policy learning. Quach *et al.* [33] combine GS with Liquid networks for real-world drone flight navigation tasks, training in simulation and then deploying the policy in the real world.

In contrast to NeRF2Real [29], which directly textures the foreground objects in simulator, we model each foreground object with editable GS parameters. Also unlike the method designed by Quach *et al.* [33], which trains drone navigation policies for target-oriented tasks, our approach involves robotic arm manipulation tasks with complex interactions.

## III. METHODS

RL-GSBridge aims to harness the potential of high-fidelity 3D GS models in Sim2Real for robot action training tasks. In

this paper, we focus on manipulation tasks for robotic arms. As shown in Fig. 1, the overall framework is divided into two parts: Real2Sim and Sim2Real. In Real2Sim stage, we collect real-world image data  $\{\mathbf{I}_k\}$ , where  $\mathbf{I}_k$  represent the image sequences of the  $k$ -th object or background in the scenarios. We will model both the geometry and appearance of the scene to build a simulation environment for the manipulation task. In Sim2Real stage, we use visual perception and deep reinforcement learning to train a policy network, and directly transfer the policy to the real world.

### A. Real2Sim: Building simulator with soft mesh binding GS

To obtain a realistic simulation environment, we use consumer-grade cameras to capture 2D image data  $\mathbf{I}_k$  of desktop-level operating platforms. Our goal is to model a geometrically accurate and texture-realistic simulation model for training operational tasks. More specifically, we use mesh models  $\{\mathbf{M}_k\}$  to represent the accurate geometric information, and Gaussian sets  $\mathbf{G}_k = \{\mathbf{g}_k^i\}$  for high-quality texture. Below, we sequentially describe the steps to construct the Simulator with GS renderer.

1) *Real-world Data Preparing*: We use a monocular camera or mobile phone to capture a 1-2 minute video of the target object on the experimental platform. We select approximately 200 keyframes from the video and use COLMAP [34] to obtain the camera's internal and external parameters for each frame. Image segmentation algorithm is also needed for extracting the target object, and we use an off-the-shelf segmenter, SAM-track [35], for efficient object segmentation.

To complete the GS model reconstruction for the simulator, it is also necessary to pre-model a geometrically accurate mesh model as a prior model for GS. Here we consider a classic and stable open-source package openMVS to obtain the corresponding mesh model.

2) *Real2Sim modeling by soft mesh binding GS*: With the object mesh, we can define Gaussian units within triangular faces as in GaMeS [9], a method that binds Gaussians onto the surface of meshes, and optimizes their properties through multi-view consistency. Vanilla GS [30] optimizes the attribute parameters of Gaussian units located at each point cloud position, where  $\theta^i = (\mu^i, r^i, s^i, \sigma^i, c^i)$  denotes the position, rotation, scale, opacity, and color of the  $i$ -th Gaussian unit  $\mathbf{g}^i$ , respectively. To achieve controllable geometric editing effects, GaMeS [9] constrains Gaussian units within the triangular mesh of the object surface, using the mean vector as the convex combination of the mesh vertices to establish the positional relationship between the Gaussian unit and the three vertices of the triangular mesh:

$$\mu^i(\alpha_1^i, \alpha_2^i, \alpha_3^i) = \alpha_1^i \mathbf{v}_1 + \alpha_2^i \mathbf{v}_2 + \alpha_3^i \mathbf{v}_3, \quad (1)$$

here,  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  represents the triangular mesh vertex positions,  $\alpha_1^i, \alpha_2^i, \alpha_3^i$  are learnable positional weight parameters for  $\mathbf{g}_i$ . However, this approach of enforced mesh-GS binding would diminish the flexibility of 3D Gaussians, limiting the optimization of Gaussian units when the mesh model is not accurate, introducing certain undesirable defects.

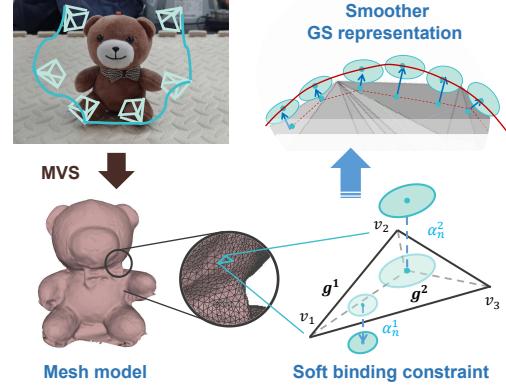


Fig. 2. Mesh-based GS Reconstruction with Soft Binding Constraints: Releasing the hard constraints of GaMeS [9] in the normal direction for smoother and more flexible object surfaces.

To address this, we propose a soft mesh binding method, which builds upon GaMeS [9], but relaxes the enforced hard mesh binding to a soft binding constraint. We introduce a component along the normal direction into the vector of positions, specifically:

$$\mu^i(\alpha_1^i, \alpha_2^i, \alpha_3^i, \alpha_n^i) = \alpha_1^i \mathbf{v}_1 + \alpha_2^i \mathbf{v}_2 + \alpha_3^i \mathbf{v}_3 + \alpha_n^i \mathbf{v}_n, \quad (2)$$

here,  $\alpha_n^i$  is a learnable weight parameter and  $\mathbf{v}_n$  is the normal vector.  $\alpha_n^i$  is constrained within the range of  $[-1, 1]$ , ensuring the association between each mesh model element and Gaussian pairs. As shown in Fig. 2, our method allows the Gaussian units within the mesh to float within a certain range along the normal vector. This flexibility in Gaussian unit optimization could bring a smoother distribution of Gaussian units on the object's surface. Ultimately, the algorithm not only ensures that the Gaussian model can still represent the accurate geometric structure according to the mesh models, but also offers some tolerance and refinement space. Besides, the binds between the mesh grid and Gaussians could even provide the possibility to handle non-rigid objects, as demonstrated in section IV-B.4.

3) *Physic Dynamics-Based GS Editing*: After obtaining the visual GS models  $\{\mathbf{G}_k\}$  and geometry models  $\{\mathbf{M}_k\}$  for real-world objects, we combine the dynamic simulation results from the simulator with real-time Gaussian model updates and render views to ensure that the visual representation follows the entire physical change process. We first use RANSAC plane regression and manual alignment methods to align GS models with mesh models in the simulator, and set the initial position of the GS model.

With the aligned initial model, we read the real-time pose changes of each object in the operational scene from the simulator. For the  $k$ -th object with its GS model  $\{\mathbf{G}_k\}$ , We acquire the rotation quaternion  $q_k$  and the homogeneous transformation matrix  $\mathbf{T}_k$  in the world coordinate system. Given Gaussian parameters  $\theta_k^i = (\mu_k^i, r_k^i, s_k^i, \sigma_k^i, c_k^i)$ , The editing process of each Gaussian  $\mathbf{g}_k^i$  in the model set  $\mathbf{G}_k$  could be formulated as:

$$\mu_k^i = \mathbf{T}_k \mu_k^i, \quad r_k^i = q_k \times r_k^i. \quad (3)$$

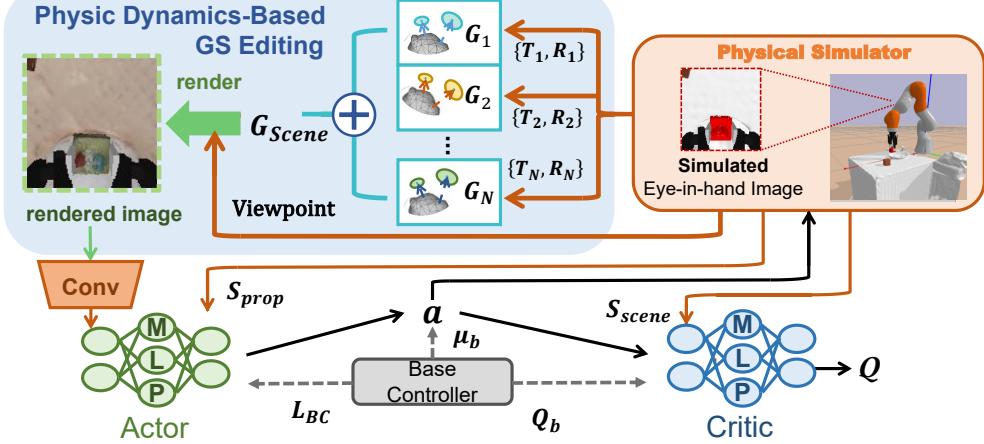


Fig. 3. Policy training pipeline in RL-GSBRidge. In the upper half of the figure, physic dynamics-based GS editing receives the transformation signals of objects and synchronizes the states of GS models. In the lower half of the figure, an actor-critic RL network receives first-person perspective images rendered by GS models as input, to learn a vision-based manipulation policy.

TABLE I  
COMPARISON OF SUCCESS RATES BETWEEN RL-GSBRIDGE AND RL-SIM IN GRASPING EXPERIMENTS, ALL CONDUCTING UNDER FOAM PAD(FP) BACKGROUND. **BOLD** INDICATES BETTER RESULTS. VALUES IN PARENTHESES REPRESENT RELATIVE CHANGE IN SUCCESS RATE DURING SIM2REAL TRANSFER. ( $\downarrow$ xx%) INDICATES A DECREASE, WHILE ( $\uparrow$ xx%) INDICATES AN INCREASE.

Object	Small_cube		Bear	
	Sim	Real	Sim	Real
RL-sim	<b>96.88</b>	12.50 ( $\downarrow$ 87%)	<b>93.75</b>	25.00 ( $\downarrow$ 73%)
RL-GSBRidge	<b>96.88</b>	<b>96.88</b>	87.50	<b>100.00</b> ( $\uparrow$ 14%)

As for local non-rigid transformations on the mesh grid, the Gaussians could be updated according to Equation 2. After applying transformations to all object models, we concat the Gaussian sets to acquire the Gaussian model of the whole scene  $\mathbf{G}_{scene}$ . Then we use rasterization to render  $\mathbf{G}_{scene}$ , obtaining the synchronized edited rendering view.

### B. Sim2Real: Train in simulation with physic dynamics-based GS renderer and zero-shot transfer to reality

We use Pybullet [36] as the simulation training platform, and employ SAC (Soft Actor-Critic) [37] algorithm for policy learning, due to its mature development in RL and widespread application in robotics. The SAC is an offline policy algorithm belonging to maximum entropy RL, which consists of two critic networks,  $Q_{\phi,1}(s, a)$  and  $Q_{\phi,2}(s, a)$ , and one actor network,  $\pi_{\theta}(s)$ . For the sampled set  $B$  from the replay buffer, the update loss of critic is defined as:

$$Loss_{critic} = \frac{1}{|B|} \sum_{(s_t, a_t, r_{t+1}, s_{t+1}) \in B} (y_t - Q_{\phi,j}(s_t, a_t))^2, \quad (4)$$

here,  $(s_t, a_t, r_{t+1}, s_{t+1})$  is a tuple belong to  $B$ , the  $y_t$  is calculated by target  $Q$  networks  $Q_{\bar{\phi},j}$ :

$$y_t = r_t + \gamma \left( \min_{j=1,2} Q_{\bar{\phi},j}(s_{t+1}, a_{t+1}) - \alpha \log \pi_{\theta}(a_{t+1} | s_{t+1}) \right). \quad (5)$$

In continuous action space, policy  $\pi_{\theta}$  outputs actions of Gaussian distribution. The loss for actor is defined as:

$$Loss_{actor} = \frac{1}{|B|} \sum_{(s_t, a_t, r_{t+1}, s_{t+1}) \in B} (\alpha \log \pi_{\theta}(\tilde{a}_t | s_t) - \min_{j=1,2} Q_{\phi,j}(s_t, \tilde{a}_t)), \quad (6)$$

where  $\tilde{a}_t$  is sampled using the reparameterization trick, i.e.,  $a_t = f_{\theta}(\epsilon_t; s_t)$ ,  $\epsilon$  is Gaussian random noise. In practical implementation, we set:  $f_{\theta} = \tanh(\mu_{\theta}(s_t) + \sigma_{\theta}(s_t) \odot \epsilon_t)$ .

Since vanilla SAC struggles to converge rapidly under sparse reward, to accelerate the learning process through automated guidance, we propose SACwB by referring to DDPGwB [38], and introduce a lightly designed baseline controller to guide the policy. This approach avoids complex reward designs while eliminating ineffective action spaces.

In SACwB, the agent executes actions from the baseline controller with probability  $\lambda$  and selects the best option between the baseline controller's and the actor's outputs with probability  $1 - \lambda$ , the objective is typically defined as:

$$y_t = r_t + \gamma \max \left( \left( \min_{j=1,2} Q_{\bar{\phi},j}(s_{t+1}, a_{t+1}) - \alpha \log \pi_{\theta}(a_{t+1} | s_{t+1}) \right), \left( \min_{j=1,2} Q_{\bar{\phi},j}(s_{t+1}, \mu_b(s_{t+1})) - \alpha \log \pi_{\theta}(\mu_b(s_{t+1}) | s_{t+1}) \right) \right). \quad (7)$$

For the supervision of the actor network, we introduce an additional behavior clone loss  $L_{bc}$  in Equation 6 to supervise the mean values of the action distribution through base controller actions, with the probability  $\lambda$ :

$$L_{bc} = \|\mu_{\theta}(s) - \mu_b(s)\|^2, \quad (8)$$

and with the probability  $1 - \lambda$ :

$$L_{bc} = \left\| \mu_{\theta}(s) - \mu \left| \arg \max \left( \min_{j=1,2} Q_{\bar{\phi},j}(s, \mu_{\theta}(s)) \right) \right. \right\|^2, \quad (9)$$

TABLE II

SIM2REAL RESULTS FOR GRASPING TASK IN VARIOUS MANIPULATION SCENARIOS. CONTENTS IN PARENTHESES AFTER THE OBJECT NAMES REPRESENT DIFFERENT BACKGROUNDS(BG), WHERE FP DENOTES FOAM PAD, AND TC DENOTES TABLECLOTH.

Object (Bg)	Cake (FP)		Banana (FP)		Small_cube (TC)		Cake (TC)		Banana (TC)		Bear (TC)	
Test scene	Sim	Real	Sim	Real	Sim	Real	Sim	Real	Sim	Real	Sim	Real
Success rate (%)	100.00	100.00	100.00	93.75 (↓6%)	96.88	87.50 (↓10%)	100.00	93.75 (↓6%)	100.00	96.88 (↓3%)	87.50	75.00 (↓14%)

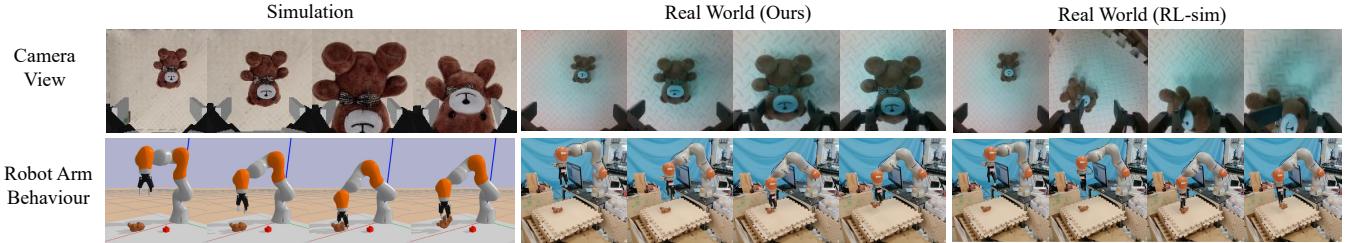


Fig. 4. Comparison of sim-to-real behavior consistency between RL-GSBridge and RL-sim.

TABLE III  
THE SIM2REAL RESULT FOR THE PICK AND PLACE TASK.

Object	Cake & Plate	
	Sim	Real
Success rate (%)	68.75	71.87 (↑4%)

here,  $\lambda$  is the decay factor, which gradually approaches 0 as training progresses.

The whole training pipeline in the simulator with physic dynamics-based GS renderer is shown in Fig. 3. The RL algorithm provides executable actions to the physics simulator, which will return state information for actor-critic learning. The GS renderer simultaneously renders the realistic images by physic dynamics-based GS rendering, serving as observation input to the actor network. We rely solely on the first-person perspective image and proprioceptive state of the robotic arm, considering several advantages as described in [39]. To enhance sim2real robustness, we add random noise to the rendered images of the GS model and randomly alter certain image attribute parameters during training for environment challenges.

After policy training in the simulator, we directly deploy the actor network onto the real robot arm, with the real-world eye-in-hand camera observations of the environment serving as visual input. The trained policy subsequently outputs the position of the end-effector and the gripper states as executable actions.

## IV. EXPERIMENTS

### A. Experiment Setup

1) *Robot Platform*: We use a KUKA iiwa robot arm paired with a Robotiq 2F-140 gripper as the platform.

The Intel RealSense D435i camera fixed on the robot arm captures the RGB images from the first-person perspective.

2) *Tasks*: As shown in Fig. 1, we design two types of tasks from the robot's first-person perspective: grasping and pick-and-place operations.

For grasping, the robot grasps a target object and lifts it up. The initial positions of the objects are randomized within a  $30 \times 30$  cm section. We use Small\_cube, Cake, Banana, and Bear as objects. As for the operation platform, we use a foam pad with and without a tablecloth as two different backgrounds. Success is considered when the object is lifted 10 cm above the table.

For pick-and-place tasks, the robotic arm pick up the cake model and place it on a plate. The position of the cake is set similarly as described in the grasping task. The plate is placed in a fixed position on the foam pad. Success is considered when the cake is placed on the plate.

3) *Evaluation Setup*: For each task, we conduct both the simulation and the real world experiments. During testing, we divide the  $30 \times 30$  cm section into four quadrants. For each task, we test the policy across a fixed number of positions in each quadrant to calculate the success rate.

4) *Baseline*: To demonstrate that our method effectively reduces Sim2Real gap, we conduct a baseline experiment called RL-sim: using the same learning method but training directly on images rendered from mesh models shaded in the PyBullet simulator. We select two representative and easily shaded objects: the Small\_cube and Bear. Grasping experiments for RL-sim are conducted on a foam pad.

### B. Experiment Results

1) *Grasping*: In Table I, we compare the grasping performance of RL strategies trained with the baseline (RL-sim) and RL-GSBridge in both simulation and real environments. For both the Small\_cube with simple geometry and textures and the Bear with complex geometry and textures, RL-sim

TABLE IV

OUR SOFT MESH BINDING METHOD COMPARED WITH GAMES [9] ON MULTIPLE FOREGROUND OBJECTS AND BACKGROUND RENDERING METRICS. SSIM IS SCALE STRUCTURAL SIMILARITY INDEX. PSNR IS PEAK SIGNAL-TO-NOISE RATIO. LPIPS IS LEARNED PERCEPTUAL IMAGE PATCH SIMILARITY. BG REFERS TO THE BACKGROUND, WHILE CONTENTS IN PARENTHESES EXPLAIN THE DETAILED DIFFERENCES.

	Banana		Bear		Cake		Small_cube		Bg (Foam Pad)		Bg (Tablecloth)		Bg (Plate)	
	Ours	GaMeS	Ours	GaMeS	Ours	GaMeS	Ours	GaMeS	Ours	GaMeS	Ours	GaMeS	Ours	GaMeS
SSIM↑	<b>0.989</b>	0.894	<b>0.964</b>	0.956	<b>0.975</b>	0.964	<b>0.989</b>	0.989	<b>0.944</b>	0.939	<b>0.781</b>	0.776	<b>0.776</b>	0.756
PSNR↑	<b>35.46</b>	26.53	<b>29.82</b>	28.18	<b>33.38</b>	30.73	<b>37.18</b>	36.64	<b>28.40</b>	27.32	<b>22.88</b>	21.86	<b>21.86</b>	20.80
LPIPS↓	<b>0.026</b>	0.049	<b>0.034</b>	0.040	<b>0.066</b>	0.079	<b>0.012</b>	0.012	<b>0.144</b>	0.149	<b>0.248</b>	0.308	<b>0.308</b>	0.311

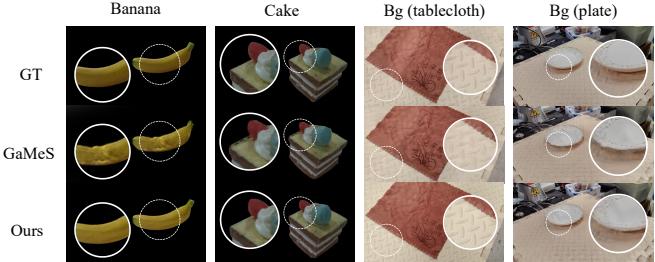


Fig. 5. Our soft binding constraint reconstruction method compared with GaMeS [9] on two foreground objects and two backgrounds.

shows a significant success rate drop when transferring to real world (an average decrease of 80%) due to visual discrepancies. In contrast, RL-GSBridge demonstrates a minor variation in success rates, maintaining high performance as in simulation. Notably, in the Bear grasping scenario, the success rate in the real world has increased by 14.28%. This may be attributed to the suboptimal simulation of Bear's soft material and non-structured shape in the simulation.

Table II shows that RL-GSBridge experiences an average drop of 6.6% of the success rate in sim-to-real transfer across various complex test scenarios, including diverse objects and desktop backgrounds. This demonstrates that the integration of the physic dynamics-based GS rendering effectively bridges the perception gap between simulation and real environments, maintaining stable strategy transfer across a range of scenarios. Additionally, we observe a noticeable decrease in success rates for the Bear in the Tablecloth background scenario, both in simulation and real environments. The primary reason for this drop is the choice of a brown tablecloth, which has similar texture features to the brown toy bear, causing difficulty in visual perception.

2) *Pick-and-Place*: As shown in Table III, we test RL-GSBridge's Sim2Real performance in pick-and-place task where a cake is placed onto a plate. The results indicate a 4.54% increase of success rate in real environments, mainly due to the differences in physical contacts between simulation and reality. In simulation, even minor excess contacts during the placement process are considered as task failures. In contrast, some contacts in real environments that do not affect the task can be tolerated, leading to better performance since the task is ultimately completed successfully.

3) *Comparison of Sim&Real Behavior Consistency*: In Fig. 4, we compare the behavior of the robotic arm in the



Fig. 6. The editing capability on non-rigid objects of our soft binding constraint GS modeling method.

simulator and in real scenarios. With the same environment, RL-GSBridge exhibits behavior highly consistent with simulation tests during manipulation, whereas RL-sim without the GS model shows significant differences. Notably, blue lights in the camera view of Fig. 4 is caused by the gripper indicator light. The image augmentation during policy training would mitigate this hue effect, ensuring the consistency of policy behavior. Meanwhile, RL-GSBridge still ensure the consistency of texture details between simulated and real-world images.

4) *GS Rendering results of different Mesh Binding approach*: In Table IV, we compare the performance of GaMeS [9] and our soft mesh binding GS model under various scenarios and achieve the SOTA performance. Furthermore, Fig. 5 shows that our method obtains fewer artifacts and more detailed texture rendering. As a supplement, in Fig. 6, our soft binding constraint GS modeling method achieves consistent results in editing a non-rigid toy bear. Unfortunately, due to the limitations in simulating soft objects in PyBullet, we do not demonstrate comprehensive experiments on deformable objects manipulation. However, this can be explored as a future direction.

## V. CONCLUSIONS

We propose RL-GSBridge, a real-to-sim-to-real framework for robotic reinforcement learning. As an attempt to apply the recently successful radiance field reconstruction methods to construct a realistic robotic simulator, RL-GSBridge has shown promising sim-to-real success rates in desktop-level tasks. This motivates us to explore future directions, such as investigating the simulation of realistic lighting [40], and integrating RL-GSBridge with advanced large-scale policy models [41] and perception learning methods [42], [43], [44]. We hope RL-GSBridge will encourage more attempts to apply radiance field reconstruction methods in robotics.

## REFERENCES

- [1] W. Zhao, J. P. Queraltà, L. Qingqing, and T. Westerlund, “Towards closing the sim-to-real gap in collaborative multi-robot deep reinforcement learning,” in *2020 5th International conference on robotics and automation engineering (ICRAE)*. IEEE, 2020, pp. 7–12.
- [2] F. Muratore, C. Eilers, M. Gienger, and J. Peters, “Bayesian domain randomization for sim-to-real transfer,” *arXiv preprint arXiv:2003.02471*, 2020.
- [3] K. Arndt, M. Hazara, A. Ghadirzadeh, and V. Kyrki, “Meta reinforcement learning for sim-to-real domain adaptation,” in *2020 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2020, pp. 2725–2731.
- [4] R. Traoré, H. Caselles-Dupré, T. Lesort, T. Sun, N. Díaz-Rodríguez, and D. Filliat, “Continual reinforcement learning deployed in real-life using policy distillation and sim2real transfer,” *arXiv preprint arXiv:1906.04452*, 2019.
- [5] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, “Nerf: Representing scenes as neural radiance fields for view synthesis,” in *ECCV*, 2020.
- [6] K. Zhang, G. Riegler, N. Snavely, and V. Koltun, “Nerf++: Analyzing and improving neural radiance fields,” *arXiv preprint arXiv:2010.07492*, 2020.
- [7] Z. Zhu, S. Peng, V. Larsson, W. Xu, H. Bao, Z. Cui, M. R. Oswald, and M. Pollefeys, “Nice-slam: Neural implicit scalable encoding for slam,” in *CVPR*, 2022, pp. 12 786–12 796.
- [8] S. Zhu, G. Wang, H. Blum, J. Liu, L. Song, M. Pollefeys, and H. Wang, “Sni-slam: Semantic neural implicit slam,” *CVPR*, 2024.
- [9] J. Waczyńska, P. Borycki, S. Tadeja, J. Tabor, and P. Spurek, “Games: Mesh-based adapting and modification of gaussian splatting,” *arXiv preprint arXiv:2402.01459*, 2024.
- [10] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, J. Pineau, et al., “An introduction to deep reinforcement learning,” *Foundations and Trends® in Machine Learning*, vol. 11, no. 3-4, pp. 219–354, 2018.
- [11] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al., “Mastering the game of go with deep neural networks and tree search,” *nature*, vol. 529, no. 7587, pp. 484–489, 2016.
- [12] Y. Deng, F. Bao, Y. Kong, Z. Ren, and Q. Dai, “Deep direct reinforcement learning for financial signal representation and trading,” *IEEE transactions on neural networks and learning systems*, vol. 28, no. 3, pp. 653–664, 2016.
- [13] X. Pan, Y. You, Z. Wang, and C. Lu, “Virtual to real reinforcement learning for autonomous driving,” *arXiv preprint arXiv:1704.03952*, 2017.
- [14] L. Pinto, M. Andrychowicz, P. Welinder, W. Zaremba, and P. Abbeel, “Asymmetric actor critic for image-based robot learning,” *arXiv preprint arXiv:1710.06542*, 2017.
- [15] Y. Liu, W. Chen, Y. Bai, J. Luo, X. Song, K. Jiang, Z. Li, G. Zhao, J. Lin, G. Li, et al., “Aligning cyber space with physical world: A comprehensive survey on embodied ai,” *arXiv preprint arXiv:2407.06886*, 2024.
- [16] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, “Domain randomization for transferring deep neural networks from simulation to the real world,” in *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2017, pp. 23–30.
- [17] K. Bousmalis, A. Irpan, P. Wohlhart, Y. Bai, M. Kelcey, M. Kalakrishnan, L. Downs, J. Ibarz, P. Pastor, K. Konolige, et al., “Using simulation and domain adaptation to improve efficiency of deep robotic grasping,” in *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2018, pp. 4243–4250.
- [18] X. Fang, A. Easwaran, B. Genest, and P. N. Suganthan, “Your data is not perfect: Towards cross-domain out-of-distribution detection in class-imbalanced data,” *ESWA*, 2024.
- [19] J. X. Wang, Z. Kurth-Nelson, D. Tirumala, H. Soyer, J. Z. Leibo, R. Munos, C. Blundell, D. Kumaran, and M. Botvinick, “Learning to reinforcement learn,” *arXiv preprint arXiv:1611.05763*, 2016.
- [20] A. A. Rusu, S. G. Colmenarejo, C. Gulcehre, G. Desjardins, J. Kirkpatrick, R. Pascanu, V. Mnih, K. Kavukcuoglu, and R. Hadsell, “Policy distillation,” *arXiv preprint arXiv:1511.06295*, 2015.
- [21] S. Zhu, G. Wang, H. Blum, J. Liu, L. Song, M. Pollefeys, and H. Wang, “Sni-slam: Semantic neural implicit slam,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 21 167–21 177.
- [22] S. Zhu, R. Qin, G. Wang, J. Liu, and H. Wang, “Semgauss-slam: Dense semantic gaussian splatting slam,” *arXiv preprint arXiv:2403.07494*, 2024.
- [23] Z. Yu, S. Peng, M. Niemeyer, T. Sattler, and A. Geiger, “Monosdf: Exploring monocular geometric cues for neural implicit surface reconstruction,” *NeurIPS*, pp. 25 018–25 032, 2022.
- [24] S. Zhi, T. Laidlow, S. Leutenegger, and A. J. Davison, “In-place scene labelling and understanding with implicit scene representation,” in *ICCV*, 2021, pp. 15 838–15 847.
- [25] M. Adamkiewicz, T. Chen, A. Caccavale, R. Gardner, P. Culbertson, J. Bohg, and M. Schwager, “Vision-only robot navigation in a neural radiance world,” *RA-L*, pp. 4606–4613, 2022.
- [26] Q. Dai, Y. Zhu, Y. Geng, C. Ruan, J. Zhang, and H. Wang, “Graspnerf: multiview-based 6-dof grasp detection for transparent and specular objects using generalizable nerf,” in *ICRA*, 2023, pp. 1757–1763.
- [27] D. Driess, I. Schubert, P. Florence, Y. Li, and M. Toussaint, “Reinforcement learning with neural radiance fields,” *NeurIPS*, 2022.
- [28] Y. Li, S. Li, V. Sitzmann, P. Agrawal, and A. Torralba, “3d neural scene representations for visuomotor control,” in *CoRL*, 2022, pp. 112–123.
- [29] A. Byravan, J. Humplík, L. Hasenclever, A. Brussee, F. Nori, T. Haarnoja, B. Moran, S. Bohez, F. Sadeghi, B. Vujatovic, et al., “Nerf2real: Sim2real transfer of vision-guided bipedal motion skills using neural radiance fields,” in *ICRA*, 2023, pp. 9362–9369.
- [30] B. Kerbl, G. Kopanas, T. Leimkuehler, and G. Drettakis, “3d gaussian splatting for real-time radiance field rendering,” *ACM Trans. Graph.*, vol. 42, no. 4, 2023.
- [31] S. Zhu, G. Wang, D. Kong, and H. Wang, “3d gaussian splatting in robotics: A survey,” *arXiv preprint arXiv:2410.12262*, 2024.
- [32] G. Lu, S. Zhang, Z. Wang, C. Liu, J. Lu, and Y. Tang, “Manigaussian: Dynamic gaussian splatting for multi-task robotic manipulation,” *arXiv preprint arXiv:2403.08321*, 2024.
- [33] A. Quach, M. Chahine, A. Amini, R. Hasani, and D. Rus, “Gaussian splatting to real world flight navigation transfer with liquid networks,” *arXiv preprint arXiv:2406.15149*, 2024.
- [34] J. L. Schönberger and J.-M. Frahm, “Structure-from-motion revisited,” in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [35] Y. Cheng, L. Li, Y. Xu, X. Li, Z. Yang, W. Wang, and Y. Yang, “Segment and track anything,” *arXiv preprint arXiv:2305.06558*, 2023.
- [36] E. Coumans and Y. Bai, “Pybullet, a python module for physics simulation for games, robotics and machine learning,” 2016.
- [37] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor,” in *International conference on machine learning*. PMLR, 2018, pp. 1861–1870.
- [38] G. Wang, M. Xin, W. Wu, Z. Liu, and H. Wang, “Learning of long-horizon sparse-reward robotic manipulator tasks with base controllers,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 3, pp. 4072–4081, 2022.
- [39] K. Hsu, M. J. Kim, R. Rafailov, J. Wu, and C. Finn, “Vision-based manipulators need to also see from their hands,” in *International Conference on Learning Representations*, 2021.
- [40] J. Gao, C. Gu, Y. Lin, Z. Li, H. Zhu, X. Cao, L. Zhang, and Y. Yao, “Relightable 3d gaussians: Realistic point cloud relighting with brdf decomposition and ray tracing,” in *European Conference on Computer Vision*. Springer, 2024, pp. 73–89.
- [41] Y. Ma, Z. Song, Y. Zhuang, J. Hao, and I. King, “A survey on vision-language-action models for embodied ai,” *arXiv preprint arXiv:2405.14093*, 2024.
- [42] R. Mendonça, S. Bahl, and D. Pathak, “Structured world models from human videos,” *arXiv preprint arXiv:2308.10901*, 2023.
- [43] X. Fang, D. Liu, P. Zhou, and G. Nan, “You can ground earlier than see: An effective and efficient pipeline for temporal sentence grounding in compressed videos,” in *CVPR*, 2023.
- [44] X. Fang, Z. Xiong, W. Fang, X. Qu, C. Chen, J. Dong, K. Tang, P. Zhou, Y. Cheng, and D. Liu, “Rethinking weakly-supervised video temporal grounding from a game perspective,” in *ECCV*, 2025.