Geometry-Aware Demonstration Augmentation for Scalable Robotic Manipulation

Michael Sha[1], Yipu Chen[2], Minghao Guo[3], Yunsheng Tian[3], Chuang Gan[4], and Wojciech Matusik[3]

Abstract— **Generalizing robot behaviours across object geometries remains a core challenge in manipulation. We propose a framework for geometry-aware demonstration augmentation that enables robust policy learning under shape variation. Starting from a single input mesh, our fully automatic geometry-augmentation pipeline produces a rich spectrum of shape variants by applying controlled stretching, compression, and local bulging while strictly preserving contact surfaces and support regions, so each instance remains structurally sound and manipulation-ready. Crucially, the augmentation process intrinsically guarantees a dense point-wise correspondence between the original and deformed geometries, allowing direct transfer of demonstration trajectories to every new shape. This yields a scalable foundation for extensive, grounded datasets of augmented demonstrations without additional human effort. We validate the pipeline on a diverse suite of household and industrial objects, generating varied shape augmentations and replaying pick-and-place demonstrations.**

I. Introduction

Robotic manipulation in unstructured environments requires policies that can generalize across a wide range of object geometries and task variations. While recent advances in behavior cloning have enabled robots to acquire complex behaviors from demonstrations ([1], [2]), these policies often struggle to transfer when object shapes deviate even slightly from the training set. This limits the scalability of learned policies, especially in real-world settings where objects exhibit diverse geometric properties, such as mugs with varying heights and widths, or assembly parts with subtle dimensional differences.

One promising direction to address this generalization gap is to leverage demonstration augmentation ([3], [4], [5], [6]) to synthetically generate variations of the original demonstrations to cover a broader distribution of object instances. Current augmentation pipelines mainly fail into two categories: **1)** Lightweight geometry transformations, such as uniform scaling, axis-aligned stretching, random bounding-box warps, and light vertex noise, are trivial to script and keep local grasp features intact yet barely perturb the underlying shape manifold and thus provide limited coverage of real-world object diversity; **2)** Fully generative pipelines, i.e., image-to-3D and text-to-3D generation, offer the opposite extreme by offering unbounded mesh varieties. However, their outputs often suffer from poor geometry quality, demand extensive post-processing, and, most critically, require a fresh demonstration for every synthesized instance before they can be used for policy learning. The labor cost of collecting or adapting demonstrations overwhelms the benefit of an unlimited mesh supply.

Neither strategy can augment demonstrations *at scale*: lightweight geometric transformations yield insufficient geometric diversity, whereas fully image- or text-based generative models incur prohibitive demonstration overhead. This exposes an intrinsic trade-off between *geometric diversity* and *demonstration efficiency*, leaving the geometry-generalization gap unresolved.

To tackle this bottleneck, we leverage recent advances in high-fidelity shape deformation and introduce a framework of geometry-aware demonstration augmentation for robotic manipulation learning. Starting from a single expert trajectory on a canonical object, we generate a family of plausible shape variants with a *slippage-preserving* deformation operator that applies controlled stretching and compression while maintaining contact surfaces and support regions. Since the deformation is bijective, it furnishes a dense point-wise mapping between the original and every altered mesh, allowing the demonstration to be transferred to each variant with negligible cost. We validate the pipeline on a representative suite of household and industrial items—including mugs, plates, brackets, screws, and gears—by generating diverse shape augmentations for each object and replaying the corresponding pick-and-place demonstrations.

II. Method

A. Slippage-Preserving Shape Deformation

Our method draws on recent advances in nonlinear shape deformation from shape analysis [7]. Unlike rigid, uniform, or axis-aligned transforms, these deformations can substantially enlarge the space of augmented shapes while guaranteeing a vertex-level correspondence between each original and warped mesh.

When deforming 3D objects, a core requirement is to preserve essential surface properties and minimize distortion. *Slippage-preserving* deformation meets this need by maintaining the local surface characteristics that enable rigid "slipping" motions. Formally, a surface is *slippable* with respect to a rigid motion when the motion translates or rotates the surface along itself without introducing empty space or overlaps. As shown in Fig. 1, this property is especially relevant in objects with *cylindrical* regions (slippable through translation along, and rotation about, the cylinder's axis), *spherical* regions (slippable via rotations around any axis through the center), and *planar* regions (slippable by

[1]Stanford Online High School
[2]Georgia Institute of Technology
[3]MIT CSAIL
[4]UMass Amherst

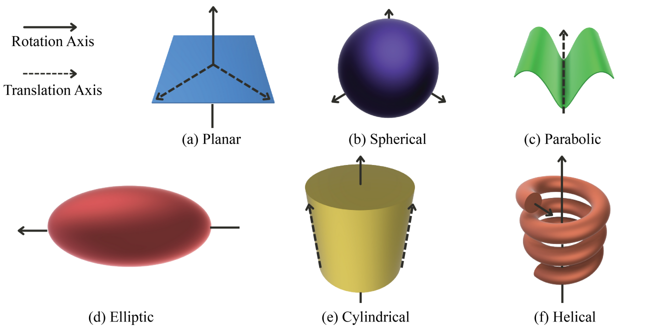

Fig. 1. Common slippable surface types. Each shape permits translation or rotation along specific axes without causing gaps or overlaps.

translations and rotations around the plane's normal). Retaining these slippable freedoms is critical for applications that depend on geometric fidelity in surface contact, such as a mug's rim thickness in robotic grasping or the snug fit of assembled parts.

Slippage preservation is particularly critical in robotic applications involving human-made objects, where contact-rich interactions and geometric fidelity directly impact task success. Everyday items such as mugs, plates, and mechanical parts often feature planar, cylindrical, or spherical surfaces that must maintain consistent thicknesses or attachment points for reliable gripping, placing, or assembly. Subtle deviations in object geometry, e.g., an elongated handle or uneven rim, can invalidate well-tuned motions, requiring extensive demonstration data or manual re-optimization. By ensuring that each object's slippable regions remain free to undergo small rigid motions (translations or rotations) even after deformation, robots can readily transfer learned trajectories to novel but structurally similar objects without losing the physical plausibility essential for robust manipulation.

B. Automatic Slippage-Preserving Deformation for Shape Augmentation

In this section, we describe a pipeline that automatically applies slippage-preserving deformation to 3D shapes to generate augmented geometry with controlled variability. Our method leverages a *cylinder-fitting* procedure to define deformation handles and displacement vectors for each object, ensuring that parts of the shape that must remain rigidly slippable can be deformed without losing the geometric properties that support robotic manipulation.

The input of our system is a triangular mesh $\mathcal{M}$ representing a human-made object of interest. The output is a deformed mesh $\mathcal{M}'$, in which key regions (such as cylindrical handles or rims) have been reshaped via slippage-preserving deformation. The method preserves crucial geometric correspondences between $\mathcal{M}$ and $\mathcal{M}'$, allowing direct transfer of robotic demonstrations (trajectories, grasps, etc.) without additional manual intervention.

Our first step is to fit a simple cylinder $\mathrm{Cyl}(\mathbf{c}, \mathbf{a}, r, h)$, where $\mathbf{c} \in \mathbb{R}^3$ is the center of the cylinder, $\mathbf{a} \in \mathbb{S}^2$ is the cylinder's axis (a unit vector), $r \in \mathbb{R}_+$ is the radius of the cylinder, and $h \in \mathbb{R}_+$ is the height of the cylinder, to the

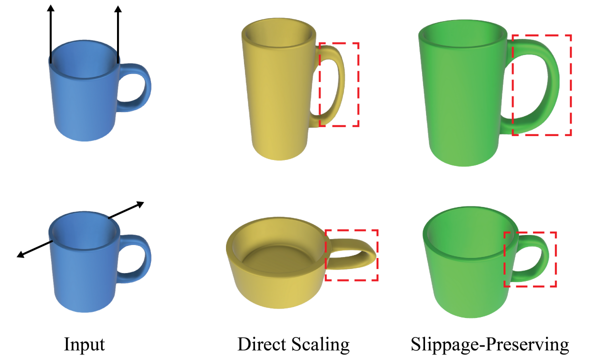

Fig. 2. Comparison between direct scaling and slippage-preserving deformation. While direct scaling may distort critical features (e.g., the shape of the handle), slippage-preserving deformation maintains the geometric fidelity of functional parts during shape modifications.

mesh region in question. To do so, we minimize an energy combining a distance term and a center-of-mass regularizer:

$$\min_{\mathbf{c}, \mathbf{a}, r, h} \sum_{\mathbf{p} \in \Omega} \mathrm{dist}\big(\mathbf{p}, \mathrm{Cyl}(\mathbf{c}, \mathbf{a}, r, h)\big) \; + \; \lambda \|\mathbf{c} - \mathbf{c}_m\|^2, \quad (1)$$

where $\Omega \subset \mathcal{M}$ is the subset of vertices belonging to the approximately cylindrical component we want to deform. The first term measures the orthogonal distance of $\mathbf{p}$ to the surface of that cylinder. The second term $\lambda \|\mathbf{c} - \mathbf{c}_m\|^2$ encourages the cylinder center $\mathbf{c}$ to remain close to the object's overall center of mass $\mathbf{c}_m$. System 1 defines a quadratic energy, which we minimize using Newton's method efficiently.

Once the cylinder parameters $(\mathbf{c}, \mathbf{a}, r, h)$ have been estimated, we specify *control handles* for the slippage-preserving deformation. These handles are the vertices that will be displaced to produce the desired new geometry. Two intuitive scenarios are included:

- **Radial Enlargement:** We place handles on the circular cross-sections at the top and bottom of the cylinder, allowing the radius r to be scaled outward or inward.
- **Height Adjustment:** We position handles along the axis $\mathbf{a}$ so that translating these points effectively "pulls" the cylinder taller or compresses it.

Consider a mug as an example. Radial enlargement lets the rim or handle grow thicker or thinner while keeping its cylindrical form intact. Height adjustment elongates or compresses the mug along its axis, producing a taller or shorter cup without altering the diameter of the opening.

For each handle vertex $\mathbf{v}_i \in \mathbb{R}^3$, we define a target position $\mathbf{v}_i^*$ that aligns with the chosen design parameters. For instance, to enlarge the radius from r to $r' > r$, we shift each handle vertex radially outward in the plane perpendicular to $\mathbf{a}$:

$$\mathbf{v}_i^* = \mathbf{c} + (\mathbf{v}_i - \mathbf{c})_\perp \left(\frac{r'}{r}\right) + (\mathbf{v}_i - \mathbf{c})_\parallel,$$

where $(\mathbf{v}_i - \mathbf{c})_\perp$ is the component of $(\mathbf{v}_i - \mathbf{c})$ perpendicular to the axis $\mathbf{a}$, and $(\mathbf{v}_i - \mathbf{c})_\parallel$ is the parallel component. Similar expressions govern height adjustments or combined changes. Additional constraints, such as preserving the wall thickness

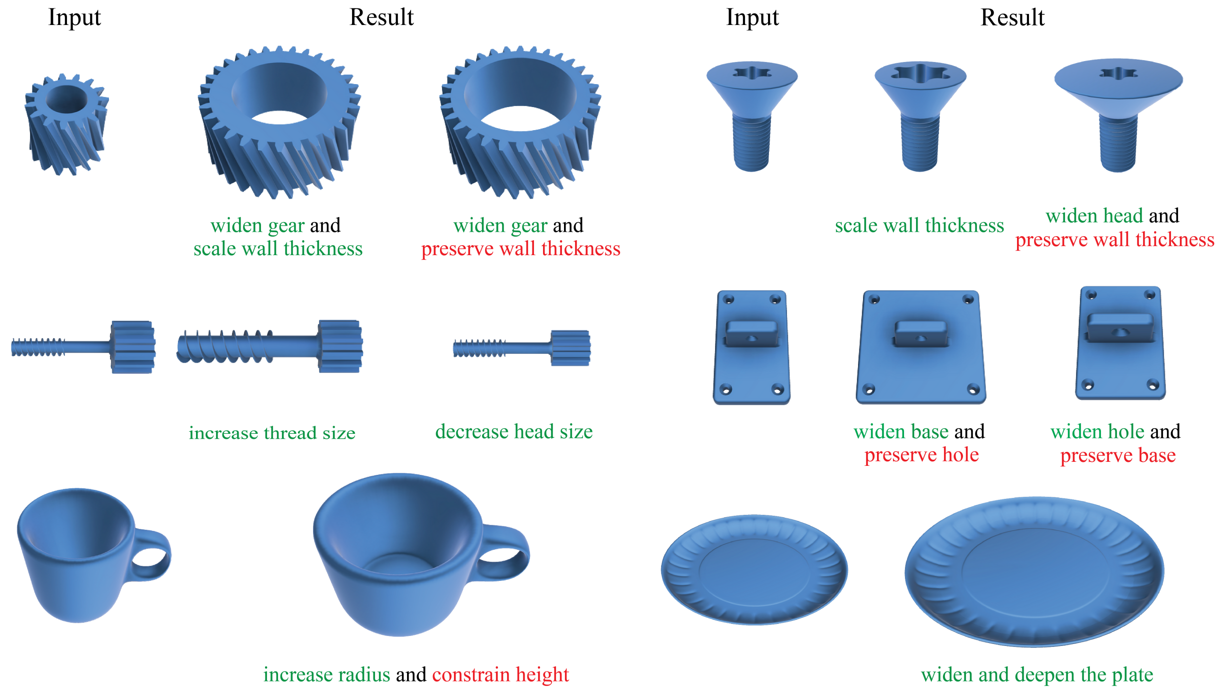

Fig. 3. Geometry-aware, slippage-preserving deformations across six object categories. For each canonical input (left) two augmented variants are shown (right), illustrating selective edits such as widening gears while either scaling or preserving wall thickness, resizing screw threads or heads, adjusting bracket bases versus holes, and independently controlling container radius, height, or depth. Green text marks attributes that are augmented, red text marks attributes that are preserved.

of a mug or preserving the hole diameter of a bracket, can also be incorporated into the deformation by constraining displacement vectors to be zero, i.e., $\mathbf{v}_i^* = \mathbf{v}_i$.

With handles and displacement vectors defined, we apply a slippage-preserving solver [7] to deform the mesh. The solver alternates between optimizing vertex positions and local transformation matrices until it converges upon a new mesh $\mathcal{M}'$ that satisfies the handle displacements to retain slippable surface motions. As shown in Fig. 2, these deformations offer far greater flexibility than simple geometry transformations such as axis-aligned scaling. Our method can preserve the general form of a mug's handle while altering specific dimensions such as radius and height. In contrast, direct scaling might distort the handle's thickness or size in undesirable ways.

C. Trajectory Transfer to Augmented Objects

The goal of trajectory transfer is to replay the original pick-and-place motion on every deformed object without re-recording demonstrations. Given a single demonstration on the canonical mesh $\mathcal{M}$, we automatically retarget both the object-centric motion and the robot's joint trajectory to every deformed variant $\mathcal{M}'$. This procedure relies only on the dense, bijective correspondence delivered by our slippage-preserving deformation.

Stage 1: Warping the object path. The original demonstration defines a sequence of object poses $\tau_{\mathcal{M}} = \{T_i\}_{i=0}^{N-1}$ from start $p^s_{\mathcal{M}}$ to goal $p^g_{\mathcal{M}}$. We express each pose as an offset $\Delta_i = S_{\mathcal{M}}(s_i)^{-1} T_i$, where $S_{\mathcal{M}}(s)$ is the unit-speed geodesic interpolating between $p^s_{\mathcal{M}}$ and $p^g_{\mathcal{M}}$ and $s_i = i/(N-1)$. For the deformed mesh we form an analogous geodesic $S_{\mathcal{M}'}(s)$ between $(p^s_{\mathcal{M}'}, p^g_{\mathcal{M}'})$ and rebuild the trajectory by $T'_i = S_{\mathcal{M}'}(s_i)\,\Delta_i$. This preserves the temporal cadence and fine-scale motion of the original path while aligning its endpoints with the geometry of $\mathcal{M}'$.

Stage 2: Transferring the grasp and joint path. At each time step i the demonstration provides an end-effector pose g_i and its contact vertex $c_i \in \mathcal{M}$. The correspondence map $\Phi : \mathcal{M} \to \mathcal{M}'$ yields the partner vertex $c'_i = \Phi(c_i)$ and guarantees consistent surface normals. We maintain the object-relative gripper transform $T_{g,i} = T_i^{-1} g_i$ by setting the new gripper pose $g'_i = T'_i T_{g,i}$. Inverse kinematics (with joint-limit and collision checks) converts $\{g'_i\}$ into a joint-space trajectory $\{q_i\}$.

Since both stages depend only on Φ and closed-form calculations in SE(3), trajectory retargeting incurs negligible computation and requires no additional optimisation or manual tuning, yet preserves grasp alignment and contact semantics across all augmented shapes.

III. Experiments

A. Implementation Details

We use four manipulation-relevant object classes: gear, fastener, bracket, and container. Each mesh is repaired as watertight and normalized to the unit bounding box. For cylinder fitting, Newton iterations stop when the gradient norm drops below 10^{-5}. The average total number of iterations is 15. For each object, we augment it with 24 sets of

TABLE I

QUANTITATIVE RESULTS ON TRAJECTORY TRANSFER.

Object Type	Average Success Rate (%)
Gear	91.7
Fastener	87.5
Bracket	95.8
Container	87.5

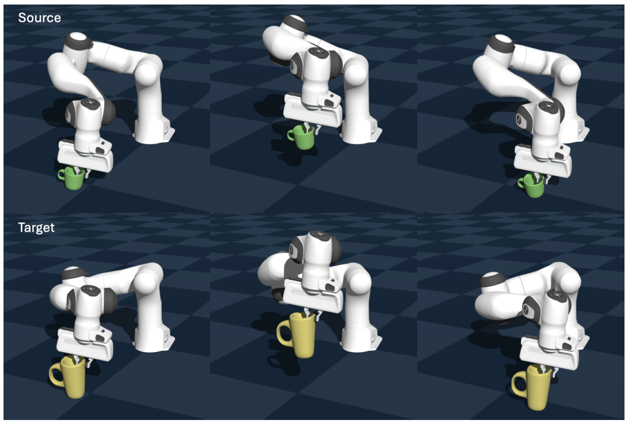

Fig. 4. Robot trajectory transfer on mug rearrangement (grasp → lift → place). Top: original demonstration on the canonical mug. Bottom: retargeted trajectory on a slippage-preserving mug variant.

handles and displacement handles as described in Sec. II-B. A single pick-and-place demonstration is recorded for every canonical mesh in Genesis [8] with a Franka Emika Panda robot.

B. Results on Shape Augmentation

Figure 3 shows qualitative results of the augmented shapes:

- **Gears**: the body can be *widened* while either *scaling* or *preserving* tooth wall-thickness;
- **Fasteners**: screw heads or threads resize independently; countersunk screws retain shank diameters;
- **Brackets**: users may *widen the base* while keeping mounting holes fixed—or conversely enlarge holes without altering the base;
- **Containers**: mugs acquire larger radii with height constraints; plates are *widened* and *deepened* without rim distortion.

C. Results on Trajectory Transfer

A transferred trajectory is defined as successful if the gripper closes with force-closure on the mapped contacts, lifts the object 20cm without slip, and places it within 2cm of the target pose.

Figure 4 and Figure 5 show representative pick-and-place executions on heavily deformed instances from two object categories: mugs and brackets. Despite substantial geometric changes, the robot successfully reuses the original demonstration to perform accurate and stable placements. Table I provides detailed statistics across all object classes, reporting average success rates over the full range of shape augmentations. While the results demonstrate that our slippage-preserving correspondence enables reliable, zero-tuning trajectory transfer even under significant shape variation, common failure cases remain. These include: (1) grasp slippage during pickup, lifting, or placement, caused by discrepancies in local curvature between the deformed and original shapes; and (2) post-transfer gripper collisions with the object, due to unexpected changes in the geometry.

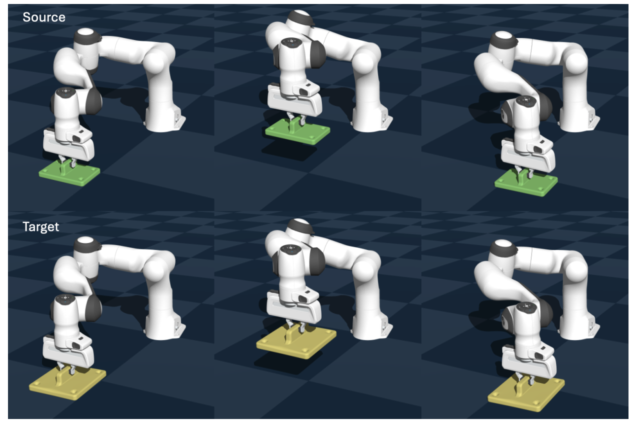

Fig. 5. Robot trajectory transfer on bracket rearrangement (grasp → lift → place). Top: original demonstration on the canonical bracket. Bottom: retargeted trajectory on a slippage-preserving bracket variant.

IV. CONCLUSION AND FUTURE WORK

We introduced an automatic, geometry-aware augmentation pipeline that generates large families of slippage-preserving shape variants from a single canonical mesh and demonstration. Future work will extend the approach to compliant and articulated objects, explore closed-loop simulation for contact-aware trajectory refinement, and investigate integration with visual perception to support end-to-end policy training and deployment in real-world settings.

REFERENCES

[1] C. Chi, S. Feng, Y. Du, Z. Xu, E. Cousineau, B. Burchfiel, and S. Song, “Diffusion policy: Visuomotor policy learning via action diffusion,” in *Proceedings of Robotics: Science and Systems (RSS)*, 2023.

[2] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, “Learning fine-grained bimanual manipulation with low-cost hardware,” *arXiv preprint arXiv:2304.13705*, 2023.

[3] T. Yu, T. Xiao, A. Stone, J. Tompson, A. Brohan, S. Wang, J. Singh, C. Tan, D. M, J. Peralta, B. Ichter, K. Hausman, and F. Xia, “Scaling robot learning with semantically imagined experience,” in *arXiv preprint arXiv:2302.11550*, 2023.

[4] E. Teoh, S. Patidar, X. Ma, and S. James, “Green screen augmentation enables scene generalisation in robotic manipulation,” *arXiv preprint arXiv:2407.07868*, 2024.

[5] A. Mandlekar, S. Nasiriany, B. Wen, I. Akinola, Y. Narang, L. Fan, Y. Zhu, and D. Fox, “Mimicgen: A data generation system for scalable robot learning using human demonstrations,” in *7th Annual Conference on Robot Learning*, 2023.

[6] J. Wang, Y. Qin, K. Kuang, Y. Korkmaz, A. Gurumoorthy, H. Su, and X. Wang, “Cyberdemo: Augmenting simulated human demonstration for real-world dexterous manipulation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 17 952–17 963.

[7] C. Araújo, N. Vining, S. Burla, M. Ruivo de Oliveira, E. Rosales, and A. Sheffer, "Slippage-preserving reshaping of human-made 3d content," *ACM Transaction on Graphics*, vol. 42, no. 6, 2023.

[8] G. Authors, "Genesis: A universal and generative physics engine for robotics and beyond," December 2024. [Online]. Available: https://github.com/Genesis-Embodied-AI/Genesis