



Goal

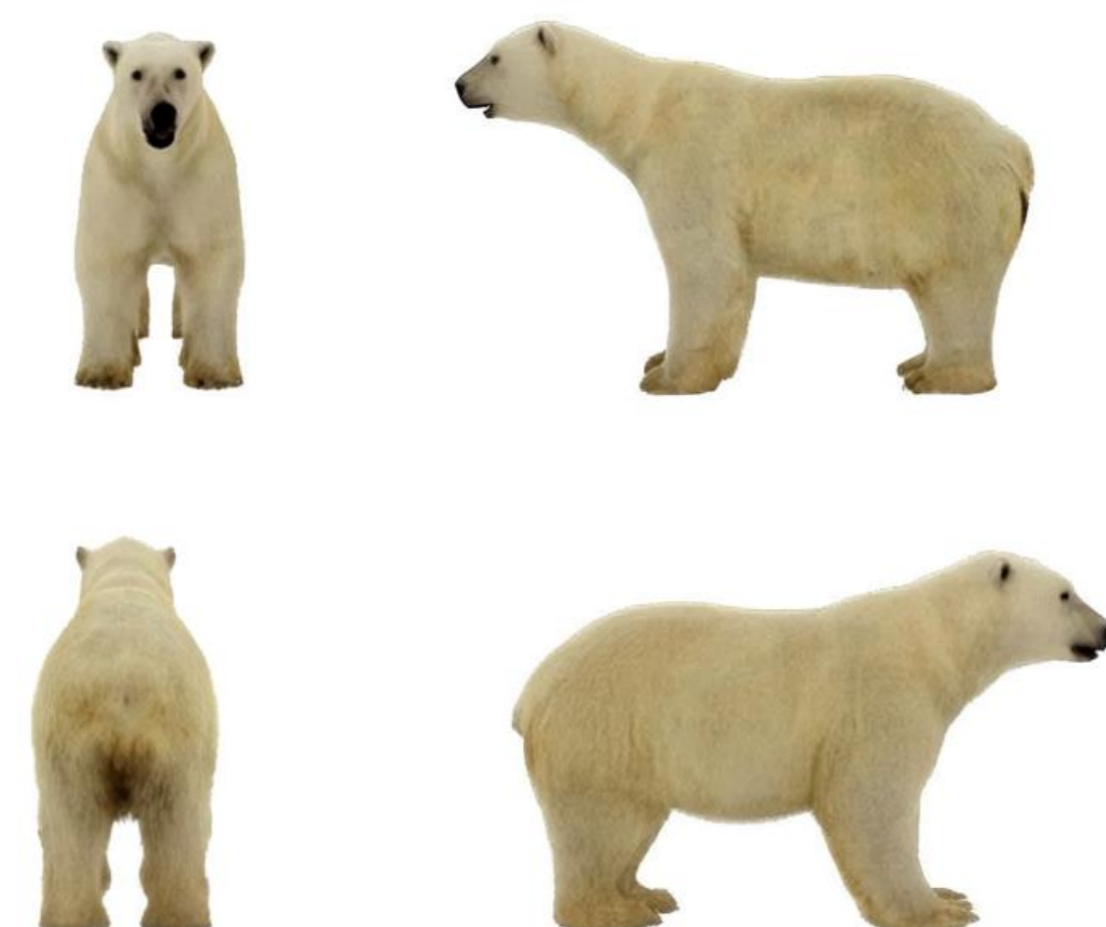
- ✓ **Goal:** One Wild Image to an Animatable Animal Avatar
- ✓ **Input:** A single in-the-wild animal image with arbitrary pose, viewpoint, occlusion, and background.
- ✓ **Output:** A visually faithful 3D animal representation that can be re-posed and rendered from novel views.
- ✓ **Challenge:** The input observes only one side, while animation requires complete appearance and topology-consistent structure.

Key Idea

- ✓ **Key Idea :** Role separation!
Canonicalize First, Reconstruct Animatable Gaussians



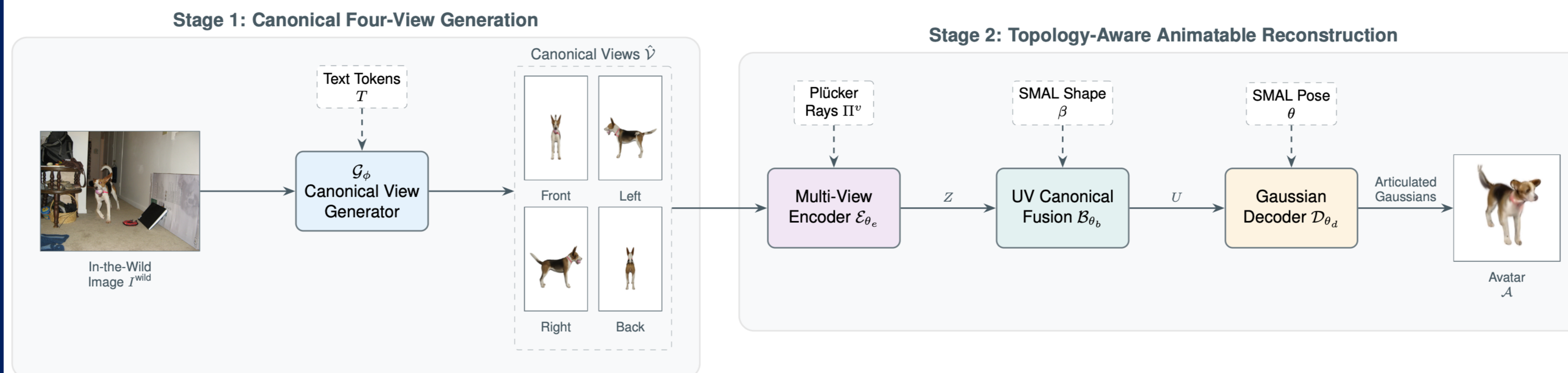
<Hard input>



<Canonical output>

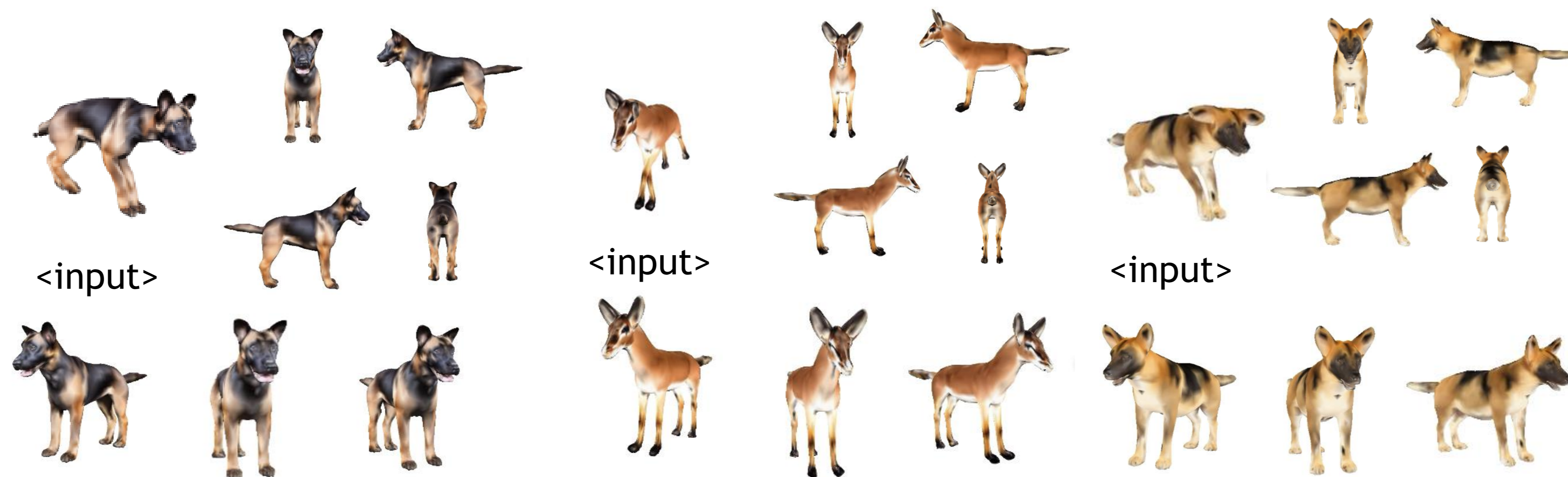
- ✓ **Stage 1: View Canonicalization**
Hard input → clean canonical four views (arbitrary pose/view/bg → canonical front/left/right/back)
- ✓ **Stage 2: Animatable Reconstruction**
Canonical four views → controllable asset (topology-aware Gaussians anchored to SMAL)

Overall Framework



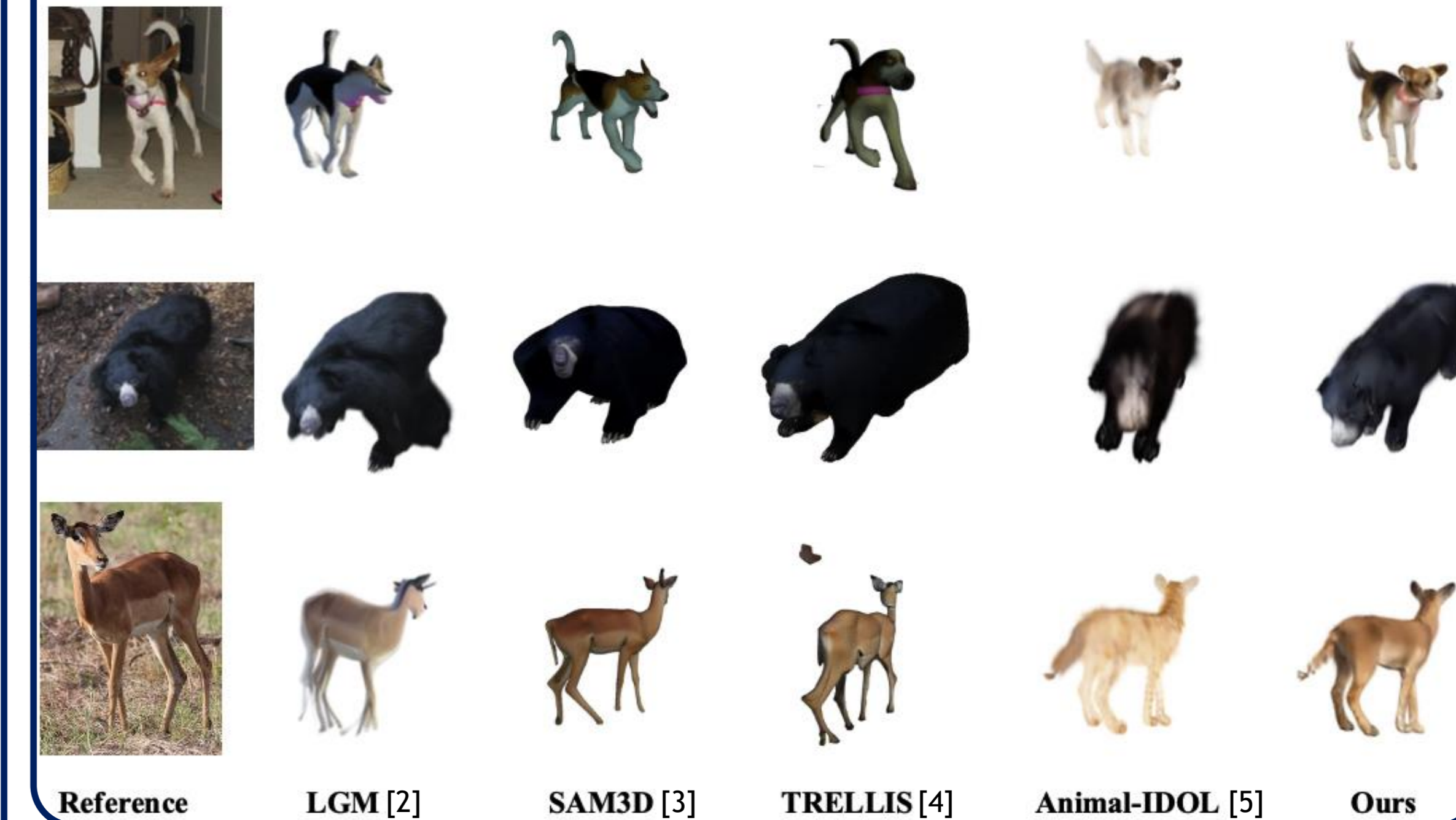
- ✓ **Stage 1: Canonical Four-View Generation**
We LoRA-train Qwen-Image-Edit^[1] to generate front / left / right / back canonical views from a single wild image, conditioned on view-specific text tokens.
- ✓ **Stage 2: Topology-Aware Animatable Reconstruction**
Canonical view features are fused in UV canonical space and decoded into SMAL-anchored Gaussians, enabling feed-forward re-posable animal reconstruction.

AniGe10K dataset

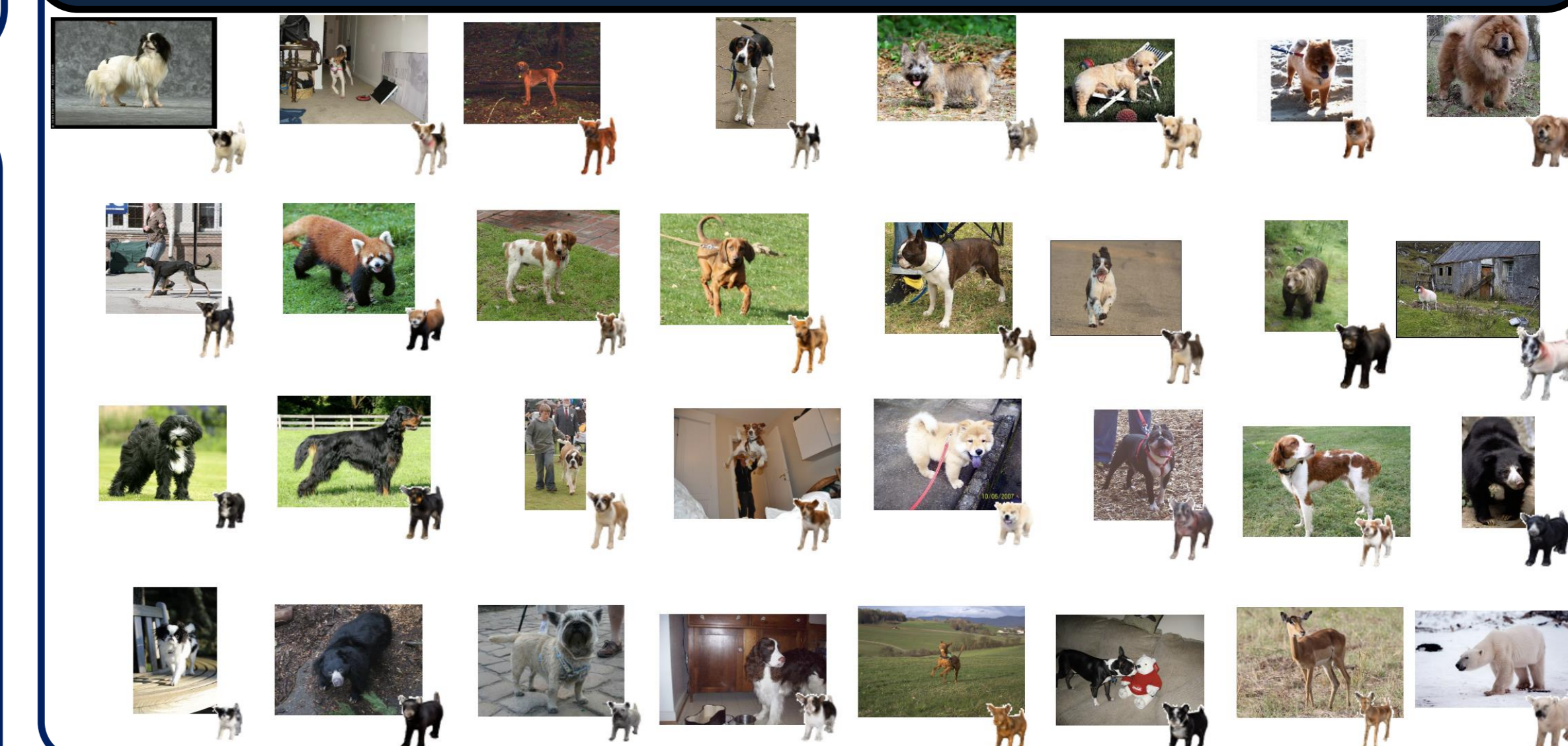


- AniGe10K each identity provides:
- ✓ Wild-style monocular input
 - ✓ 4 canonical views: front / left / right / back (for stage1)
 - ✓ 24 canonical renderings for full-view supervision (for stage2)
 - ✓ Exact SMAL shape, pose, camera, and topology annotations

Qualitative Result



Additional Result



Future work

- ✓ Improve data quality and diversity.
- ✓ Enhance canonical view generator.
- ✓ Propose a benchmark for single-image animatable animal reconstruction.

[1] Wu, Chenfei, et al. "Qwen-Image Technical Report". arXiv 2025.
 [2] Tang, Jiaxiang, et al. "LGM: Large Multi-View Gaussian Model for High-Resolution 3D Content Creation". ECCV 2024.
 [3] Chen, Xingyu, et al. "SAM 3D: 3Dfy Anything in Images". CVPR 2025.
 [4] Xiang, Jianfeng, et al. "TRELIS: Structured 3D Latents for Scalable and Versatile 3D Generation". CVPR 2025.
 [5] Zhuang, Yiyu, et al. "IDOL: Instant Photorealistic 3D Human Creation from a Single Image". CVPR 2025.