

Steering 3D Generations: Preference Alignment via Direct Reward and Preference Optimization

Yuanhang Wang¹, Lizhe Qi^{1†}, and Wenqiang Zhang^{1†}

College of Intelligent Robotics and Advanced Manufacturing, Fudan University,
Shanghai, China

yuanhangwang23@m.fudan.edu.cn

Abstract. Generating high-fidelity and controllable 3D assets from a single image remains a significant challenge. We propose a novel framework that integrates preference optimization into a rectified flow-based diffusion model operating on a latent set representation. This approach enhances geometric fidelity and user alignment, enabling the efficient synthesis of diverse 3D formats. Our core contribution is a versatile reward-based optimization strategy for aligning generation with human-centric criteria like surface quality. We introduce Direct Reward Optimization (DRO), which fine-tunes the model using only absolute quality feedback, removing the need for pairwise data. The framework also seamlessly incorporates Direct Preference Optimization (DPO) when paired data is available. Parameter-efficient control is achieved via Low-Rank Adaptation (LoRA). To ensure data quality, we also develop a robust preprocessing pipeline using a neural SDF reconstruction method with Gaussian curvature constraints to produce high-quality watertight meshes. Extensive experiments show our framework achieves state-of-the-art performance. Quantitative evaluations demonstrate significant improvements in Chamfer Distance and F-Score, while qualitative results confirm superior geometric fidelity and user preference alignment compared to existing methods.

Keywords: 3D Generation · Preference Alignment · Direct Reward Optimization · Direct Preference Optimization

1 Introduction

The creation of digital 3D content is undergoing a paradigm shift, driven by the rapid advancements in AI-generated content (AIGC). As demand for high-quality 3D assets explodes across industries ranging from entertainment and gaming to industrial design and the metaverse, the need for automated, efficient, and controllable generation methods has become more pressing than ever. Despite remarkable progress, a substantial quality and utility gap persists between 3D generative models and their 2D counterparts [47, 50]. This disparity stems from fundamental challenges unique to the 3D domain: the inherent complexity

[†] Co-Corresponding Authors.



Fig. 1: Our method generates high-quality, controllable 3D assets from a single image in seconds. By integrating preference optimization (DPO/DRO) into a rectified flow-based diffusion model, our framework significantly enhances geometric fidelity and alignment with user-centric criteria, producing visually appealing and structurally sound models suitable for diverse applications from creative design to engineering.

of geometric representations (e.g., meshes, point clouds, neural fields) [25,38] and the chronic scarcity of large-scale, high-quality 3D datasets [9]. Consequently, generating 3D assets that are not only geometrically accurate but also semantically aligned with user intent remains a significant open challenge.

Recent diffusion-based models [56, 66, 68] have emerged as the dominant paradigm for single-image-to-3D generation, demonstrating impressive capabilities. However, a closer examination reveals two fundamental limitations that hinder their practical applicability. First, their geometric quality is often compromised. Trained on vast but noisy web-crawled data, these models frequently produce assets plagued with topological errors, non-watertight surfaces, floating artifacts, and inconsistent surface normals.

Second, and perhaps more critically, aligning the generated outputs with nuanced, user-centric preferences is exceptionally difficult. Beyond simple text-to-shape correspondence, users often have implicit quality criteria, such as desiring a "smooth" car body, "sharp" mechanical edges, or a certain "aesthetic appeal." These preferences are subjective, context-dependent, and notoriously difficult to formalize into a differentiable loss function. The current generation process often acts as a "black box," lacking robust mechanisms for fine-grained control and optimization beyond the initial conditioning prompt. This lack of alignment severely limits the utility of generated assets in professional workflows that demand precision and artistic control.

To overcome these limitations, we argue that a paradigm shift is required, moving beyond simple statistical learning towards explicit preference-aware optimization. We propose that a two-pronged strategy is essential: one must first establish a foundation of high-quality data to ensure geometric plausibility, and then explicitly align the generative model with human-centric quality criteria. This paper introduces a novel framework that operationalizes this strategy. To address the data challenge, we propose a robust data curation pipeline that

leverages a neural Signed Distance Function (SDF) reconstruction method. This allows us to produce a high-quality, watertight training dataset with controllable geometric features, forming a solid basis for our model.

More critically, to bridge the alignment gap, we introduce a versatile reward-based optimization strategy that fine-tunes the generative model based on explicit preference signals. Our key innovation is the adaptation of powerful preference optimization techniques from the domain of large language models to the structured, geometric domain of 3D generation. We introduce two complementary approaches: Direct Preference Optimization (DPO) [45] for scenarios with paired preference data, and our novel Direct Reward Optimization (DRO), which learns from absolute quality feedback without requiring pairwise comparisons. This allows our model to learn and internalize what constitutes a "good" 3D model from a human perspective.

Our entire framework is built upon a powerful rectified flow-based diffusion model, which provides a stable and efficient generative backbone. To enable precise and efficient control over the generation, we employ Low-Rank Adaptation (LoRA) [22] during the preference-tuning stage. Our main contributions can be summarized as follows:

- **Reward-Based Preference Optimization:** We introduce a versatile optimization strategy featuring Direct Reward Optimization (DRO) for absolute feedback and Direct Preference Optimization (DPO) for paired data, effectively aligning 3D generation with human preferences for geometric quality.
- **Parameter-Efficient Control:** We employ Low-Rank Adaptation (LoRA) to achieve efficient, multi-modal controllable generation (e.g., via text or images) with minimal computational overhead, preserving the base model’s generalization capabilities.
- **High-Quality Data Curation:** We develop a robust data pipeline using a neural SDF with Gaussian curvature constraints to produce high-quality, watertight training meshes with controllable features from raw, unoriented point clouds.

Extensive experiments demonstrate that our framework achieves state-of-the-art performance, significantly improving geometric fidelity and user preference alignment. We believe our work provides a valuable blueprint for building the next generation of controllable and high-quality 3D generative models, paving the way for their broader adoption in real-world applications, as shown in Figure 1.

2 Related Work

Our work is situated at the intersection of large-scale 3D generative modeling and preference-based model alignment. We review the most relevant literature in these areas.

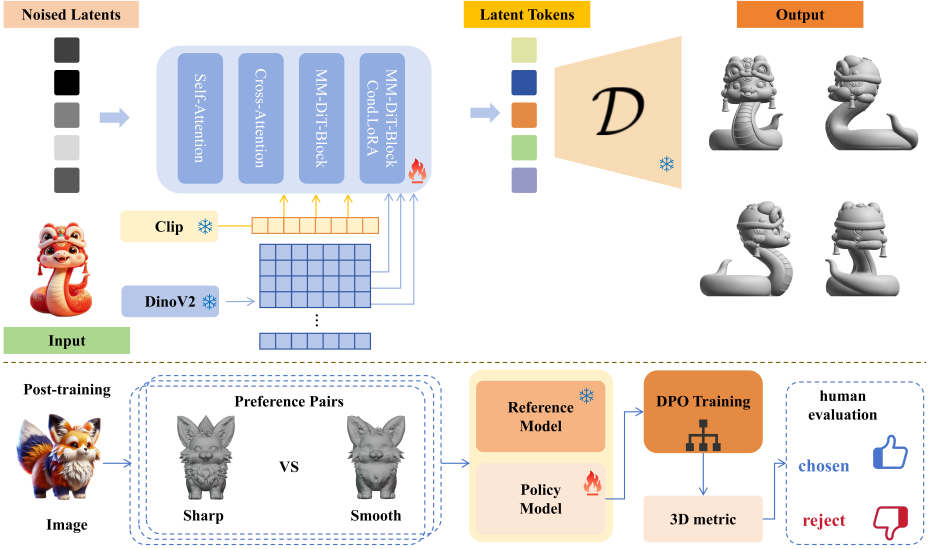


Fig. 2: An overview of our pipeline. It consists of two main stages: (i) a robust data curation pipeline that generates high-quality, watertight meshes with controllable geometric features using a neural SDF reconstruction method; and (ii) a preference-driven, single-image-to-3D generation model. The model, built upon a rectified flow-based diffusion architecture, is fine-tuned using our proposed reward optimization strategies (DRO/DPO) and parameter-efficient techniques (LoRA) to align with user preferences and achieve controllable, high-fidelity output.

Large-scale 3D Datasets. The availability of large-scale datasets is fundamental to the success of generative models [14, 19, 24, 46, 47]. While the 2D and language domains benefit from massive data corpora [2, 11, 13, 44], the 3D domain has historically been data-constrained. Early datasets like ShapeNet [3] and ABO [7] provided foundational geometric shapes but lacked the scale and diversity needed for training highly generalizable models [54]. The recent emergence of extensive collections, notably Objaverse [9] and its successor Objaverse-XL [8], has been a turning point. These datasets, containing millions of objects, have enabled systematic investigation of scaling laws in 3D synthesis [6, 8]. However, as these datasets are largely web-crawled, they often suffer from significant quality issues, including non-manifold geometry, inconsistent topology, and missing textures. This dirty data poses a major challenge for training models that can produce high-fidelity outputs. Our work directly addresses this by proposing a robust data curation pipeline to filter and enhance publicly available data.

3D Generation from 2D Models. Leveraging powerful pre-trained 2D diffusion models and robust visual representations [17, 18, 26, 34], many recent works synthesize 3D content without direct 3D supervision. This line of research was pioneered by DreamFusion [43], which introduced Score Distillation Sampling (SDS)

to optimize a Neural Radiance Field (NeRF) [38]. However, SDS is known to suffer from issues like low diversity and overly saturated results. To mitigate this, subsequent works proposed improvements such as Variational Score Distillation (VSD) in ProlificDreamer [58]. Another line of work focused on improving the 3D representation itself for better fidelity and efficiency. This includes explicit mesh representations, such as the tetrahedral meshes in Magic3D [30] or the deformable meshes in Fantasia3D [4]. The advent of 3D Gaussian Splatting [25] led to highly efficient methods like DreamGaussian [55] and GaussianDreamer [64]. A third popular paradigm is to first generate consistent multi-view images and then perform neural reconstruction. This includes view-set diffusion models like Zero-1-to-3 [32] and MVDream [52], which are often paired with powerful reconstruction backbones like InstantMesh [62], LRM [21], or geometry bridges like Hi3DGen [63]. While these 2D-lifted methods are highly flexible, they often struggle with geometric consistency (the "Janus" multi-face problem) and overall quality, as the 2D priors do not always guarantee a coherent 3D structure. Other concurrent efforts have also further addressed fast and consistent texture generation to complement these geometries [1]. This fundamental limitation motivates the exploration of methods trained directly on 3D data.

Direct 3D Generative Models. Directly training on 3D data offers the potential for superior geometric fidelity and consistency. Early efforts using Generative Adversarial Networks (GANs) [10, 16, 51] or autoregressive models [39] demonstrated feasibility but faced scalability issues. The field has since shifted towards diffusion models, which have shown greater stability and quality. Initial diffusion-based approaches operated on explicit 3D representations. This includes models for point clouds [37, 40], which are simple but lack surface information, and models for voxel grids [69], which suffer from high memory consumption and limited resolution. To overcome these limitations, a dominant trend is to operate in a compressed latent space, a strategy proven highly effective in image generation [47]. In 3D, this involves training an autoencoder to compress shapes into a compact latent code, and then training a diffusion model in this latent space. Various latent representations have been explored, such as explicit voxel latents in Shape-E [23], implicit function-based latents in SDFusion [6], and unstructured latent sets in 3DShape2VecSet [65]. Further benchmarking and sampling refinements for 3D shape VAEs have also been proposed by Dora [5]. Recent state-of-the-art models like CLAY [66], Trellis [61], and Direct3D [60] have further refined this latent diffusion approach, achieving remarkable generation quality, while models like Unleashing VecSet [27] and Craftsman3D [29] have driven fast shape generation and native high-fidelity mesh refinement. Despite this progress, these models are typically trained with simple, fixed loss functions (e.g., Chamfer Distance or L2 loss on latents) that do not fully capture human perception of quality. Consequently, there is a lack of mechanisms to explicitly align the generation process with user-centric criteria like surface smoothness or feature sharpness.

Aligning Generative Models with Human Preferences. In the domain of large language models (LLMs), aligning model outputs with human intent has been a major breakthrough. This was largely driven by techniques like Reinforcement Learning from Human Feedback (RLHF) [41], which involves training a reward model on human preference data and then fine-tuning the language model using reinforcement learning. Furthermore, principles from causal inference [35, 42, 49] and knowledge distillation [15, 20, 48] have increasingly inspired robust knowledge transfer and generative alignment architectures. While powerful, RLHF can be complex and unstable to implement. More recently, Direct Preference Optimization (DPO) [45] has emerged as a more stable and direct method for preference alignment, bypassing the need for an explicit reward model by leveraging a clever loss function derived from the reward-policy relationship. These techniques have been instrumental in creating powerful and helpful models like ChatGPT. Inspired by this success, preference alignment has also been explored in image generation [57], where it has been used to improve aesthetics and prompt-following. To date, however, such explicit preference alignment has been largely unexplored in the 3D generation domain, where defining and optimizing for "quality" is arguably even more complex and multifaceted work bridges this gap. We build upon a direct 3D latent diffusion model and introduce a preference optimization framework tailored for 3D. By adapting DPO and proposing our own variant, DRO (which is particularly suited for scenarios without paired preference data), we explicitly fine-tune the model to satisfy geometric and aesthetic criteria that are important to human users but difficult to formalize with traditional loss functions.

3 Methodology

Our methodology introduces a comprehensive framework for generating high-fidelity, controllable 3D assets from a single image. This framework is designed to address the critical limitations of existing methods in geometric accuracy and alignment with user-centric quality criteria. This section first outlines the foundational architecture of our generative model. Subsequently, we detail our three core contributions: a robust data curation pipeline for producing superior training data, a novel preference alignment strategy using reward optimization, and a parameter-efficient approach for achieving fine-grained controllable generation. The pipeline architecture of the entire model can be referred to as Figure2. The data processing procedure can be referred to as follows as Figure4.

3.1 Rectified Flow for 3D Generation

Our approach builds upon a state-of-the-art latent set diffusion framework, inspired by recent advancements like Hunyuan3D-2 [56], which has demonstrated strong performance in shape synthesis. The architecture consists of two main components: a Shape VAE and a rectified flow-based diffusion model. The Shape

Double-Trough Prior as a Geometric Selector

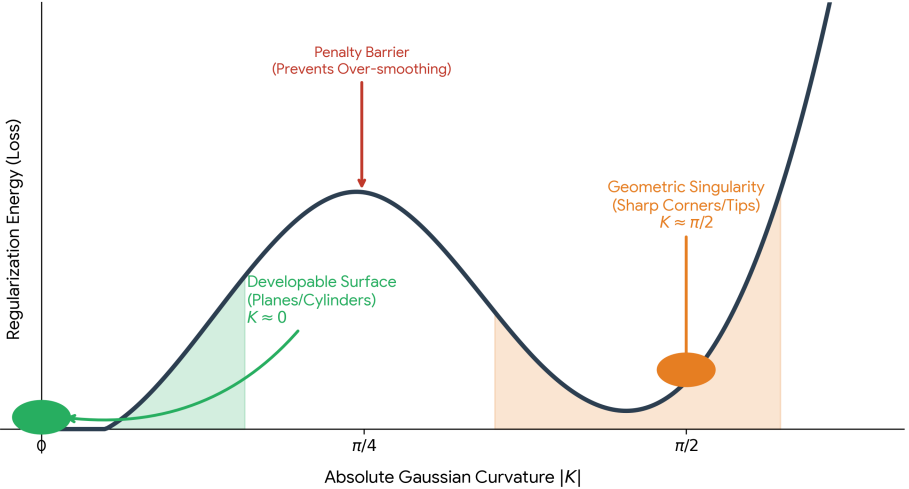


Fig. 3: The Double-Trough Geometric Prior as an Energy Landscape. Unlike standard regularization that indiscriminately minimizes curvature ($|K| \rightarrow 0$), our prior acts as a geometric selector with two distinct basins of attraction: (1) The Developable Basin (Green, $|K| \approx 0$) encourages smooth, developable patches typical of mesh surfaces; (2) The Singularity Basin (Orange, $|K| \approx \pi/2$) preserves sharp topological features like corners and ridges. Crucially, a Penalty Barrier (Red) separates these states, penalizing the ambiguous, over-smoothed curvatures often associated with noise. This effectively "snaps" the generated geometry into a clean, piecewise developable manifold.

VAE first encodes a 3D mesh into a compact, unordered latent set representation, which is adept at handling the variable topology and complexity inherent in 3D data. The diffusion model then learns to generate this latent set conditioned on an input image.

We adopt the Rectified Flow [33] paradigm, a highly efficient flow-matching objective that offers distinct advantages over traditional score-based diffusion models. By defining a linear probability path between noise and data, it simplifies the learning objective to predicting a constant velocity field. This results in straighter and more stable generation trajectories, enabling faster and higher-quality sampling with fewer steps. The optimization objective is formally expressed as:

$$\mathcal{L}_{\text{flow}} = \mathbb{E}_{t, \mathbf{x}_0, \mathbf{x}_1} [\|\mathbf{u}_\theta(\mathbf{x}_t, \mathbf{c}, t) - (\mathbf{x}_1 - \mathbf{x}_0)\|_2^2] \quad (1)$$

where $\mathbf{c}$ represents the conditioning information (e.g., image embeddings from CLIP). This robust and efficient generative foundation is ideal for our subsequent preference-based fine-tuning, as it provides a stable baseline upon which nuanced adjustments can be made.

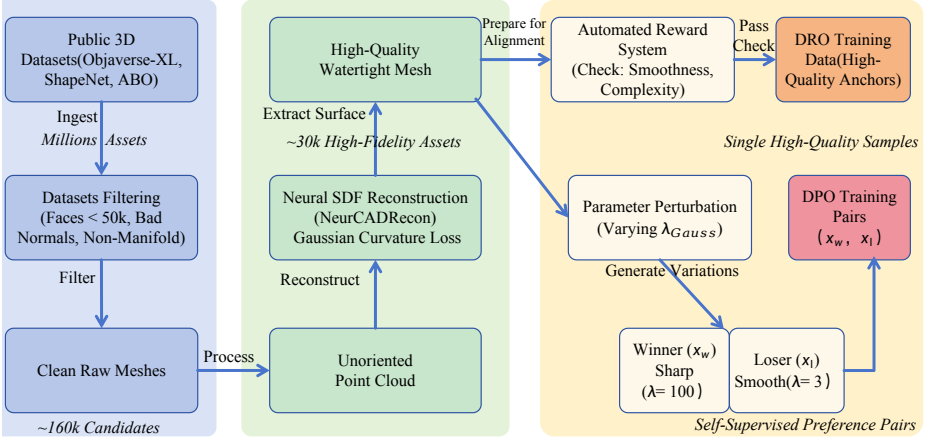


Fig. 4: Data Curation Pipeline. We distill millions of raw assets into high-quality training data through a three-stage process: (1) Heuristic filtering of raw inputs; (2) Neural geometry refinement using NeurCADRecon to ensure watertightness; and (3) Automated generation of single-point anchors for DRO and preference pairs for DPO via curvature parameter perturbation.

3.2 High-Quality Data Curation via Neural SDFs

The quality of training data is paramount for high-fidelity 3D generation. Large-scale public datasets like Objaverse [9], while vast, suffer from significant quality issues, including non-manifold geometry, holes, and inconsistent topology. Training on such noisy data directly compromises the geometric integrity of the generated assets. To address this, we establish a robust data preprocessing pipeline to create a high-quality, curated dataset.

A critical step in this pipeline is the reconstruction of clean, watertight meshes from raw point clouds. For this task, we employ the neural SDF reconstruction method proposed in NeurCADRecon [12]. We specifically chose this method for its unique capability to enforce geometric priors relevant to our goals. Its core innovation is a geometrically-aware loss function that penalizes non-zero Gaussian curvature, encouraging the generation of piecewise-developable surfaces characteristic of many manufactured objects. The geometric loss term is defined as:

$$\mathcal{L}_{\text{Gauss}} = \frac{1}{|\Omega|} \int_{\Omega} \text{DT}(|k_{\text{Gauss}}(\mathbf{x})|) d\mathbf{x}, \quad (2)$$

where k_{Gauss} is the Gaussian curvature and $\text{DT}(\cdot)$ is a double-trough function that strategically preserves sharp features. A key advantage for our framework is that by adjusting the weight λ_{Gauss} of this loss term, we can controllably generate a continuous spectrum of surface styles, from perfectly smooth to distinctly sharp-featured. This capability is not only crucial for creating a diverse, high-quality training dataset but is also instrumental in systematically generating

the preference pairs (smooth vs. sharp) required for our reward optimization strategies, as detailed in which can be referred to as shown in Figure3.

3.3 Preference Alignment via Reward Optimization

Our primary contribution is a versatile reward-based optimization strategy to align the generative model with human-centric preferences that are often poorly captured by standard pixel- or point-level losses like Chamfer Distance. We reformulate the alignment task as a reward optimization problem, enabling fine-tuning without costly test-time optimization. We propose two complementary objectives based on the nature of the available preference data.

The Reward Oracle for DRO. Direct Reward Optimization (DRO) requires only absolute quality feedback for each sample, provided by a reward oracle $o(\mathbf{x}_0) \in \{0, 1\}$. A key challenge is defining an oracle that is both scalable and reflective of human judgment. In our framework, we design a practical, rule-based oracle that programmatically assesses the geometric quality of a generated mesh $\mathbf{x}_0$. A sample is deemed "preferred" ($o(\mathbf{x}_0) = 1$) if it satisfies a set of predefined criteria:

- **Watertightness:** The mesh must be a closed, manifold surface with no holes. This is a fundamental indicator of a valid 3D model and is checked using standard topological algorithms.
- **Sufficient Complexity:** The face count must exceed a minimum threshold. This simple yet effective heuristic prevents the model from collapsing to trivial or degenerate shapes.
- **Surface Smoothness:** The variance of the vertex-wise mean curvature across the mesh must be below a certain value. High variance indicates noise, artifacts, or other undesirable surface irregularities. This metric serves as a robust proxy for perceptual surface quality.

This automated, multi-faceted oracle provides a scalable and consistent way to generate reward signals for fine-tuning without any manual annotation.

Direct Reward Optimization (DRO). Given the binary feedback from the oracle, we formulate the DRO objective to maximize the likelihood of preferred samples and minimize it for non-preferred ones. This is achieved by re-weighting the standard diffusion objective, leading to a practical loss function:

$$\mathcal{L}_{\text{DRO}} = \mathbb{E} \left[w(t)(1 - 2o(\mathbf{x}_0)) \|\epsilon - \epsilon_{\theta}(\mathbf{x}_t, t, I)\|_2^2 \right], \quad (3)$$

where the expectation is over data, time t , and noise ϵ . The term $(1 - 2o(\mathbf{x}_0))$ acts as a dynamic learning signal modulator: it amplifies the gradient for preferred samples and suppresses or reverses it for non-preferred ones. This effectively encourages the model to denoise preferred samples more accurately, steering the entire generation process towards higher-quality regions of the solution space.

Direct Preference Optimization (DPO). When pairwise preference data $(\mathbf{x}_0^w, \mathbf{x}_0^l)$ (a preferred and a non-preferred sample) is available, we employ the standard DPO framework. Grounded in the Bradley-Terry model, DPO translates a preference ranking into a direct loss on the model’s parameters. The objective maximizes the margin between the implicit rewards of the preferred and non-preferred samples:

$$\mathcal{L}_{\text{DPO}} = -\mathbb{E} \left[\log \text{sigmoid} \left(\beta' (\Delta(\mathbf{x}_t^l) - \Delta(\mathbf{x}_t^w)) \right) \right], \quad (4)$$

where $\Delta(\mathbf{x}_t) = \|\epsilon - \epsilon_\theta(\mathbf{x}_t, t)\|_2^2 - \|\epsilon - \epsilon_{\text{ref}}(\mathbf{x}_t, t)\|_2^2$ is the difference in denoising error between the fine-tuned and reference models, and β' is a scaling factor. This loss directly encourages the denoising error for preferred samples to decrease relative to the reference model, providing a powerful signal for fine-grained alignment.

3.4 Parameter-Efficient Controllable Fine-Tuning

A significant challenge in adapting large generative models is the risk of catastrophic forgetting and the prohibitive computational cost of full fine-tuning. To enable multi-modal, controllable generation (e.g., via text or semantic labels) in a parameter-efficient manner, we integrate Low-Rank Adaptation (LoRA) [22] into our fine-tuning process.

Instead of updating the millions of parameters in the pre-trained diffusion model ϵ_θ , we freeze the original weights and inject small, trainable low-rank matrices into the transformer architecture. Specifically, we augment the query and value projection matrices in the self-attention layers with a low-rank update $\Delta W = BA$, where B and A are the trainable matrices. During fine-tuning with our DRO or DPO objectives, only these LoRA parameters are optimized. This approach allows the model to adapt to new tasks and preferences with a minimal number of trainable parameters (typically 1% of the original model). It effectively preserves the strong, diverse knowledge learned during pre-training while enabling precise, user-driven control over attributes like style and structure, thus achieving a balance between generalization and specialization.

4 Experiment

This section systematically evaluates our proposed framework. We first detail the experimental setup, including implementation details, datasets, and evaluation metrics. We then present a comprehensive comparison against state-of-the-art (SOTA) methods, including quantitative and qualitative results. Finally, we conduct in-depth ablation studies to validate our design choices and report the results of a user study on human preferences.

4.1 Experimental Setup

Implementation Details. Our framework is built upon the Hunyuan3D-2 [56] open-source model, a state-of-the-art image-to-3D generator with a rectified flow

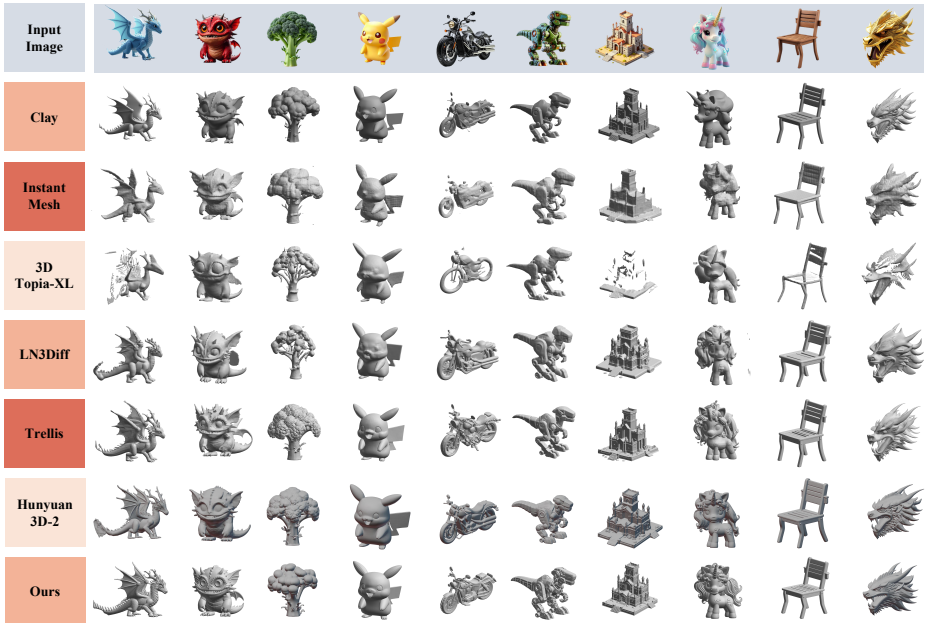


Fig. 5: Visual comparisons of generated 3D assets between our method and previous approaches.

transformer architecture. For all fine-tuning experiments, we employ LoRA [22] with a rank of 64, targeting only the linear layers of the coarse geometry transformer, as we found this sufficient for effective preference alignment. We use the AdamW [36] optimizer with a learning rate of 1×10^{-4} and a batch size of 64, training on 8 NVIDIA A100 GPUs. For classifier-free guidance, we use a strength of 2.8. The DRO model was trained for 100,000 steps, while the DPO model was trained for 200,000 steps. The β parameter in the DPO loss (Eq. 4) was set to 500. For the rule-based reward oracle in DRO, we set the minimum face count to 500,000 and the maximum mean curvature variance to 0.1, in addition to a standard watertightness check.

Datasets. Our training data curation process begins with large-scale public datasets, including ObjaverseXL and Objaverse [9], ShapeNet [3], and ABO [7]. We first apply a rigorous filtering pipeline to remove low-quality assets, discarding those with fewer than 50k faces, incorrect normals, or transparent materials. The remaining assets are then processed using the neural SDF reconstruction method described in Sec. 3.2 to create a high-quality, watertight dataset of approximately 100,000 models. For DPO training, we generate 80,000 preference pairs by varying the $\mathcal{L}_{\text{Gauss}}$ parameter during reconstruction to create smooth (non-preferred) and sharp (preferred) versions of the same object. For DRO training, we use the entire curated dataset, with rewards assigned by our auto-

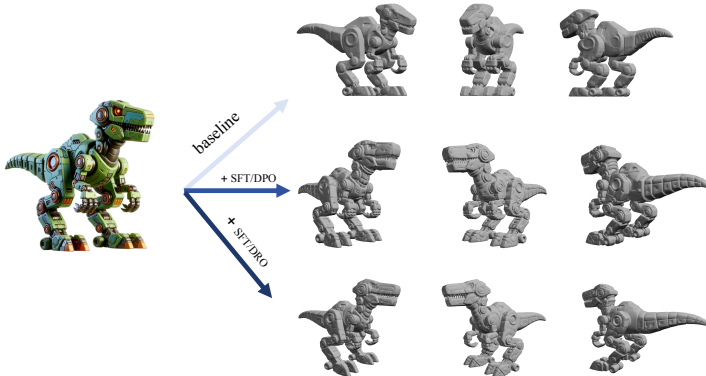


Fig. 6: The ablation study of different training methods. Starting from the baseline generator, SFT on curated data improves geometric completeness, while preference-based fine-tuning further enhances surface quality. DRO yields the cleanest geometry and the best quantitative performance.

mated oracle. For quantitative evaluation, we use the unseen Toys4k [53] dataset to ensure a fair comparison.

Metrics. We evaluate the geometric quality of the generated meshes using two standard metrics: *Chamfer Distance* (*CD*) for overall shape similarity and *F-Score* (with a threshold of 0.05) for surface detail accuracy. To evaluate the quality of rendered views, we also report *PSNR* and *LPIPS* [67] on rendered normal maps. Following established practices [31, 59], all meshes are normalized to a unit cube and aligned with the ground truth using ICP before metric computation.

Table 1: Comparison of geometric reconstruction results.

Method	CD↓	F-score↑	PSNR-N↑	LPIPS-N↓
Instantmesh	0.0224	62.78	27.34	0.096
3DTopia-XL	0.0228	65.90	31.87	0.090
LN3Diff [28]	0.0299	59.17	27.10	0.094
Clay	0.0124	72.59	35.35	0.035
Trellis	0.0083	73.12	36.11	0.024
Hunyuan3D-2	0.0089	72.79	35.53	0.033
Ours	0.0036	77.97	37.05	0.021

4.2 Comparison with State-of-the-Art Methods

Quantitative Analysis. Table 1 presents the quantitative comparison of our method against several SOTA baselines. Our framework, fine-tuned with ei-

ther DRO or DPO, demonstrates superior performance across all geometric and rendering-based metrics. Notably, our DRO-tuned model achieves a Chamfer Distance of 0.0036, a significant improvement over the trellis (0.0083) and other strong competitors like Clay (0.0124). This empirically validates our central hypothesis: explicitly optimizing for geometric quality via preference alignment leads to substantial gains in generation fidelity, even when starting from a powerful pre-trained model.

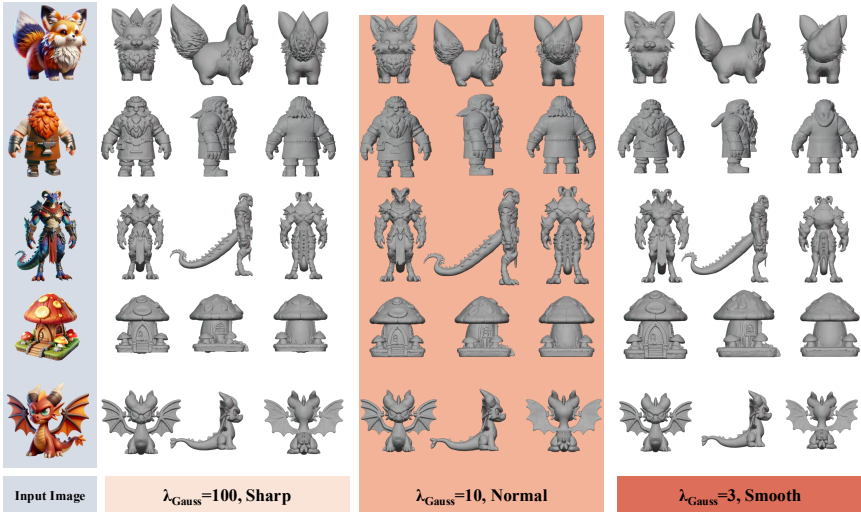


Fig. 7: The smoothness control of the reconstructed surface is achieved by controlling the Gaussian curvature.

Qualitative Analysis. Figure 5 provides a visual comparison of our results against other methods. Our generated assets exhibit significantly fewer surface artifacts and preserve sharper, more coherent details compared to competitors. For instance, where other models may produce noisy or overly smoothed surfaces, our method generates clean topology and well-defined features. Figure 7 further showcases the controllability of our framework. By leveraging the smooth and sharp labels during DPO training, we can steer the generation towards a desired surface style, demonstrating that our preference alignment mechanism is not only effective for quality improvement but also for user-driven stylistic control.

4.3 Ablation Study

We conduct ablation studies to validate the effectiveness of our proposed components. As shown in Table 2, we compare our DRO and DPO fine-tuning against the original baseline and a standard Supervised Fine-Tuning (SFT) approach on our curated high-quality dataset. All fine-tuning methods outperform the

baseline, confirming the benefit of our high-quality dataset. Interestingly, while SFT improves geometry, both DPO and DRO achieve even better results. We hypothesize that SFT, by simply training on good data, may slightly overfit to the specific geometric distributions in our curated set. DPO, which learns from relative preferences, shows stronger performance, validating the benefit of preference-based learning. Most notably, our DRO-tuned model achieves the best performance. We attribute this to DRO’s simpler and more direct learning signal. By rewarding any sample that meets the robust criteria of our automated oracle, DRO encourages the model to explore a broader region of high-quality solutions without being constrained by potentially noisy or ambiguous pairwise comparisons. This seems to lead to better generalization and overall geometric fidelity on the unseen test set, which can be referred to as Figure6.

Table 2: Ablation of supervised and preference-based fine-tuning. DRO achieves the best geometric fidelity, indicating that absolute geometric reward signals provide an effective alignment signal.

Method	Geometry	
	CD↓	F-Score↑
Baseline	0.0079	73.79
Baseline+SFT	0.0045	76.28
Baseline+SFT/DPO	0.0048	76.77
Baseline+SFT/DRO	0.0036	77.97

4.4 User Study

To assess whether the geometric improvements are perceived by human users, we conducted a two-alternative forced-choice user study with 90 participants and 45 input images. For each input image, we generated one 3D asset using Ours-DRO and one using Clay. The input image and the two generated 3D models were shown side-by-side in randomized order under the same rendering protocol, and participants were asked to select the result that better matched the input image in terms of geometric fidelity, surface cleanliness, and overall structural plausibility. Across all collected pairwise comparisons, Ours-DRO received 0.6 of the votes, indicating a consistent perceptual preference over Clay in this study.

5 Conclusion

In this work, we presented a preference-aligned framework for high-fidelity 3D generation. By synergizing a geometrically-aware data curation pipeline with our novel Direct Reward Optimization (DRO) and DPO strategies, we effectively steer a rectified flow model towards superior structural integrity and user intent. Extensive experiments demonstrate state-of-the-art performance in generating clean, watertight assets. Future avenues include learning aesthetic reward models from human feedback and extending this paradigm to complex scene generation.

References

1. Bensadoun, R., Kleiman, Y., Azuri, I., Harosh, O., Vedaldi, A., Neverova, N., Gafni, O.: Meta 3d texturegen: Fast and consistent texture generation for 3d objects. arXiv preprint arXiv:2407.02430 (2024)
2. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Nee-lakantan, A., Shyam, P., Sastry, G., Askell, A., et al.: Language models are few-shot learners. *Advances in neural information processing systems* **33**, 1877–1901 (2020)
3. Chang, A.X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., Savarese, S., Savva, M., Song, S., Su, H., et al.: Shapenet: An information-rich 3d model repository. arXiv preprint arXiv:1512.03012 (2015)
4. Chen, R., Chen, Y., Jiao, N., Jia, K.: Fantasia3d: Disentangling geometry and appearance for high-quality text-to-3d content creation. In: *Proceedings of the IEEE/CVF international conference on computer vision*. pp. 22246–22256 (2023)
5. Chen, R., Zhang, J., Liang, Y., Luo, G., Li, W., Liu, J., Li, X., Long, X., Feng, J., Tan, P.: Dora: Sampling and benchmarking for 3d shape variational auto-encoders. In: *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR)*. pp. 16251–16261 (June 2025)
6. Cheng, Y.C., Lee, H.Y., Tulyakov, S., Schwing, A.G., Gui, L.Y.: Sdfusion: Multi-modal 3d shape completion, reconstruction, and generation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 4456–4465 (2023)
7. Collins, J., Goel, S., Deng, K., Luthra, A., Xu, L., Gundogdu, E., Zhang, X., Vicente, T.F.Y., Dideriksen, T., Arora, H., et al.: Abo: Dataset and benchmarks for real-world 3d object understanding. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 21126–21136 (2022)
8. Deitke, M., Liu, R., Wallingford, M., Ngo, H., Michel, O., Kusupati, A., Fan, A., Laforte, C., Voleti, V., Gadre, S.Y., et al.: Objaverse-xl: A universe of 10m+ 3d objects. *Advances in Neural Information Processing Systems* **36** (2024)
9. Deitke, M., Schwenk, D., Salvador, J., Weihs, L., Michel, O., Vanderbilt, E., Schmidt, L., Ehsani, K., Kembhavi, A., Farhadi, A.: Objaverse: A universe of annotated 3d objects. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 13142–13153 (2023)
10. Deng, Y., Yang, J., Xiang, J., Tong, X.: Gram: Generative radiance manifolds for 3d-aware image generation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2022)
11. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
12. Dong, Q., Xu, R., Wang, P., Chen, S., Xin, S., Jia, X., Wang, W., Tu, C.: Neurcadrecon: Neural representation for reconstructing cad surfaces by enforcing zero gaussian curvature. *ACM Transactions on Graphics* **43**(4) (July 2024). <https://doi.org/10.1145/3658171>
13. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)
14. Esser, P., Kulal, S., Blattmann, A., Entezari, R., Müller, J., Saini, H., Levi, Y., Lorenz, D., Sauer, A., Boesel, F., et al.: Scaling rectified flow transformers for high-resolution image synthesis. In: *ICML* (2024)

15. Fang, G., Song, J., Wang, X., Shen, C., Wang, X., Song, M.: Contrastive model inversion for data-free knowledge distillation. arXiv preprint arXiv:2105.08584 (2021)
16. Gao, J., Shen, T., Wang, Z., Chen, W., Yin, K., Li, D., Litany, O., Gojcic, Z., Fidler, S.: Get3d: A generative model of high quality 3d textured shapes learned from images. *Advances In Neural Information Processing Systems* **35**, 31841–31854 (2022)
17. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.: Masked autoencoders are scalable vision learners. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 16000–16009 (2022)
18. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 770–778 (2016)
19. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems* **30** (2017)
20. Hinton, G., Vinyals, O., Dean, J., et al.: Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531 **2**(7) (2015)
21. Hong, Y., Zhang, K., Gu, J., Bi, S., Zhou, Y., Liu, D., Liu, F., Sunkavalli, K., Bui, T., Tan, H.: Lrm: Large reconstruction model for single image to 3d. In: *The Eleventh International Conference on Learning Representations (ICLR)* (2024)
22. Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen., W.: Lora: Low-rank adaptation of large language models. In: *ICLR* (2022)
23. Jun, H., Nichol, A.: Shape-e: Generating conditional 3d implicit functions. arXiv preprint arXiv:2305.02463 (2023)
24. Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 4401–4410 (2019)
25. Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G.: 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.* **42**(4), 139–1 (2023)
26. Krizhevsky, A., Hinton, G., et al.: Learning multiple layers of features from tiny images (2009)
27. Lai, Z., Zhao, Y., Zhao, Z., Liu, H., Wang, F., Shi, H., Yang, X., Lin, Q., Huang, J., Liu, Y., Jiang, J., Guo, C., Yue, X.: Unleashing vecset diffusion model for fast shape generation (2025), <https://arxiv.org/abs/2503.16302>
28. Lan, Y., Hong, F., Yang, S., Zhou, S., Meng, X., Dai, B., Pan, X., Loy, C.C.: Ln3diff: Scalable latent neural fields diffusion for speedy 3d generation. In: *ECCV* (2024)
29. Li, W., Liu, J., Chen, R., Liang, Y., Chen, X., Tan, P., Long, X.: Craftsman: High-fidelity mesh generation with 3d native generation and interactive geometry refiner. arXiv preprint arXiv:2405.14979 (2024)
30. Lin, C.H., Gao, J., Tang, L., Takikawa, T., Zeng, X., Huang, X., Kreis, K., Fidler, S., Liu, M.Y., Lin, T.Y.: Magic3d: High-resolution text-to-3d content creation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 300–309 (2023)
31. Liu, M., Xu, C., Jin, H., Chen, L., Varma T, M., Xu, Z., Su, H.: One-2-3-45: Any single image to 3d mesh in 45 seconds without per-shape optimization. *NeurIPS* (2023)
32. Liu, R., Wu, R., Van Hoorick, B., Tokmakov, P., Zakharov, S., Vondrick, C.: Zero-1-to-3: Zero-shot one image to 3d object. In: *Proceedings of the IEEE/CVF international conference on computer vision*. pp. 9298–9309 (2023)

33. Liu, X., Gong, C., Liu, Q.: Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003* (2022)
34. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 10012–10022 (2021)
35. Lopez-Paz, D., Nishihara, R., Chintala, S., Scholkopf, B., Bottou, L.: Discovering causal signals in images. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 6979–6987 (2017)
36. Loshchilov, I.: Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101* (2017)
37. Luo, S., Hu, W.: Diffusion probabilistic models for 3d point cloud generation. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 2837–2845 (2021)
38. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM* **65**(1), 99–106 (2021)
39. Nash, C., Ganin, Y., Eslami, S.A., Battaglia, P.: Polygen: An autoregressive generative model of 3d meshes. In: *International conference on machine learning*. pp. 7220–7229. PMLR (2020)
40. Nichol, A., Jun, H., Dhariwal, P., Mishkin, P., Chen, M.: Point-e: A system for generating 3d point clouds from complex prompts. *arXiv preprint arXiv:2212.08751* (2022)
41. Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C.L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., Lowe, R.: Training language models to follow instructions with human feedback. In: *NeurIPS* (2022)
42. Pearl, J.: Causal inference in statistics: An overview. *Statistics surveys* **3**, 96–146 (2009)
43. Poole, B., Jain, A., Barron, J.T., Mildenhall, B.: Dreamfusion: Text-to-3d using 2d diffusion. In: *The Eleventh International Conference on Learning Representations* (2023)
44. Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., et al.: Learning transferable visual models from natural language supervision. In: *International Conference on Machine Learning*. pp. 8748–8763. PMLR (2021)
45. Rafailov, R., Sharma, A., Mitchell, E., Manning, C.D., Ermon, S., Finn, C.: Direct preference optimization: Your language model is secretly a reward model. In: *NeurIPS* (2023)
46. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., Chen, M.: Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125* (2022)
47. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 10684–10695 (2022)
48. Romero, A., Ballas, N., Kahou, S.E., Chassang, A., Gatta, C., Bengio, Y.: Fitnets: Hints for thin deep nets. *arXiv preprint arXiv:1412.6550* (2014)
49. Rubin, D.B.: Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association* **100**(469), 322–331 (2005)

50. Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E.L., Ghasemipour, K., Gontijo Lopes, R., Karagol Ayan, B., Salimans, T., et al.: Photorealistic text-to-image diffusion models with deep language understanding. In: *NeurIPS (2022)*
51. Sharma, G., Goyal, R., Liu, D., Kalogerakis, E., Maji, S.: Csgnet: Neural shape parser for constructive solid geometry. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (June 2018)*
52. Shi, Y., Wang, P., Ye, J., Long, M., Li, K., Yang, X.: Mvdream: Multi-view diffusion for 3d generation. *arXiv preprint arXiv:2308.16512 (2023)*
53. Stojanov, S., Thai, A., Rehg, J.M.: Using shape to categorize: Low-shot learning with an explicit shape bias. In: *CVPR (2021)*
54. Sun, C., Shrivastava, A., Singh, S., Gupta, A.: Revisiting unreasonable effectiveness of data in deep learning era. In: *Proceedings of the IEEE international conference on computer vision*. pp. 843–852 (2017)
55. Tang, J., Ren, J., Zhou, H., Liu, Z., Zeng, G.: Dreamgaussian: Generative gaussian splatting for efficient 3d content creation. *arXiv preprint arXiv:2309.16653 (2023)*
56. Team, T.H.: Hunyuan3d 2.0: Scaling diffusion models for high resolution textured 3d assets generation (2025)
57. Wallace, B., Dang, M., Rafailov, R., Zhou, L., Lou, A., Purushwalkam, S., Ermon, S., Xiong, C., Joty, S., Naik, N.: Diffusion model alignment using direct preference optimization. In: *CVPR (2024)*
58. Wang, Z., Lu, C., Wang, Y., Bao, F., Li, C., Su, H., Zhu, J.: Prolificdreamer: High-fidelity and diverse text-to-3d generation with variational score distillation. *Advances in Neural Information Processing Systems* **36** (2024)
59. Wang, Z., Wang, Y., Chen, Y., Xiang, C., Chen, S., Yu, D., Li, C., Su, H., Zhu, J.: Crm: Single image to 3d textured mesh with convolutional reconstruction model. In: *ECCV (2024)*
60. Wu, S., Lin, Y., Zhang, F., Zeng, Y., Xu, J., Torr, P., Cao, X., Yao, Y.: Direct3d: Scalable image-to-3d generation via 3d latent diffusion transformer. *arXiv preprint arXiv:2405.14832 (2024)*
61. Xiang, J., Lv, Z., Xu, S., Deng, Y., Wang, R., Zhang, B., Chen, D., Tong, X., Yang, J.: Structured 3d latents for scalable and versatile 3d generation. *arXiv preprint arXiv:2412.01506 (2024)*
62. Xu, J., Cheng, W., Gao, Y., Wang, X., Gao, S., Shan, Y.: Instantmesh: Efficient 3d mesh generation from a single image with sparse-view large reconstruction models. *arXiv preprint arXiv:2404.07191 (2024)*
63. Ye, C., Wu, Y., Lu, Z., Chang, J., Guo, X., Zhou, J., Zhao, H., Han, X.: Hi3dgen: High-fidelity 3d geometry generation from images via normal bridging. *arXiv preprint arXiv:2503.22236 (2025)*
64. Yi, T., Fang, J., Wang, J., Wu, G., Xie, L., Zhang, X., Liu, W., Tian, Q., Wang, X.: Gaussiandreamer: Fast generation from text to 3d gaussians by bridging 2d and 3d diffusion models. In: *CVPR (2024)*
65. Zhang, B., Tang, J., Niessner, M., Wonka, P.: 3dshape2vecset: A 3d shape representation for neural fields and generative diffusion models. *ACM Transactions on Graphics (TOG)* **42**(4), 1–16 (2023)
66. Zhang, L., Wang, Z., Zhang, Q., Qiu, Q., Pang, A., Jiang, H., Yang, W., Xu, L., Yu, J.: Clay: A controllable large-scale generative model for creating high-quality 3d assets. *ACM Transactions on Graphics (TOG)* **43**(4), 1–20 (2024)
67. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 586–595 (2018)

68. Zhao, Z., Liu, W., Chen, X., Zeng, X., Wang, R., Cheng, P., Fu, B., Chen, T., Yu, G., Gao, S.: Michelangelo: Conditional 3d shape generation based on shape-image-text aligned latent representation. *Advances in neural information processing systems* **36**, 73969–73982 (2023)
69. Zhou, L., Du, Y., Wu, J.: 3d shape generation and completion through point-voxel diffusion. In: *Proceedings of the IEEE/CVF international conference on computer vision*. pp. 5826–5835 (2021)