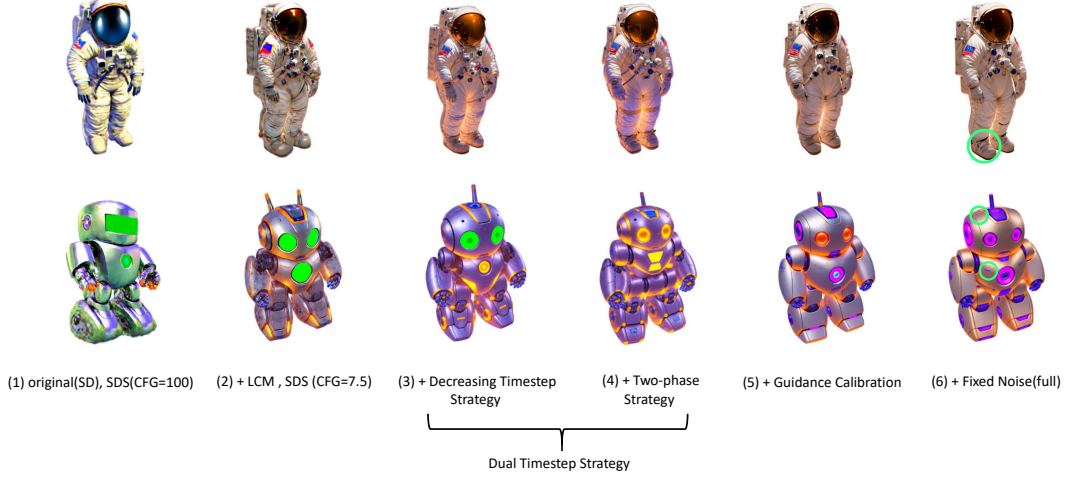


# DreamLCM: Towards High Quality Text-to-3D Generation Via Latent Consistency Model

Anonymous Authors



**Figure 1: Ablation experiments.** Comparing column (1) to column (2), a significant enhancement in generating more details is demonstrated by incorporating LCM into DreamLCM. Furthermore, compared to columns (2) and (3), an improvement in generating fine details, such as the robot’s eyes, is observed. Subsequently, when comparing column (3) to column (4), geometric improvements are evident. Furthermore, comparing column (4) to column (5), a better generation quality is observed. Finally, comparing column (5) to column (6), it can be observed that in column (6) DreamLCM focuses on generating specific objects, thereby reducing the occurrence of the "feature-average" outcome, as shown in green. For instance, while the robot in column (5) exhibits blemishes on its head and chest due to the "feature-average" phenomenon, this issue is mitigated in column (6). Furthermore, the shoes in the astronaut example are generated with improved quality.

We conduct more experiments to further illustrate the superiority of DreamLCM. Initially, we present additional ablation studies, explaining the specific contributions of each component within DreamLCM towards enhancing generation quality. Subsequently, we showcase additional generated samples, demonstrating the ability of DreamLCM to produce diverse 3D objects with fine details. Finally, we provide further comparisons with existing works, where DreamLCM generates more photo-realistic objects, thus substantiating that it achieves state-of-the-art results. The settings of the experiments are the same as in the body of the paper.

## 1 MORE ABLATION STUDIES

We conduct more ablation studies, as shown in Fig. 1.

## 2 MORE GENERATED SAMPLES

As shown in Fig 2 and 3, We present more generated samples, indicating that our works can produce high-quality 3D objects with fine details.

## 3 MORE COMPARISONS WITH BASELINES

We conduct further comparisons with current works, e.g., Dreamfusion [5] ProlificDreamer [6] and LucidDreamer [3], corresponding to the Discussion section in the body of the paper. We obtain some

results of Dreamfusion and ProlificDreamer from previous studies [2, 4], while the other results are sourced from threestudio [1]. As shown in Fig. 4, our work achieves comparable, and in some cases superior, generation quality and finer details compared to these baselines. Notably, ProlificDreamer requires approximately 10 hours of training per sample. Therefore, we illustrate that our approach is more efficient than ProlificDreamer.

## REFERENCES

- [1] Yuan-Chen Guo, Ying-Tian Liu, Ruizhi Shao, Christian Laforte, Vikram Voleti, Guan Luo, Chia-Hao Chen, Zi-Xin Zou, Chen Wang, Yan-Pei Cao, and Song-Hai Zhang. 2023. threestudio: A unified framework for 3D content generation. <https://github.com/threestudio-project/threestudio>.
- [2] Oren Katzir, Or Patashnik, Daniel Cohen-Or, and Dani Lischinski. 2023. Noise-Free Score Distillation. *arXiv:2310.17590*
- [3] Yixun Liang, Xin Yang, Jiantao Lin, Haodong Li, Xiaogang Xu, and Yingcong Chen. 2023. LucidDreamer: Towards High-Fidelity Text-to-3D Generation via Interval Score Matching. *arXiv:2311.11284*
- [4] Yuanze Lin, Ronald Clark, and Philip Torr. 2024. DreamPolisher: Towards High-Quality Text-to-3D Generation via Geometric Diffusion. *arXiv:2403.17237 [cs.CV]*
- [5] Ben Poole, Ajay Jain, Jonathan T. Barron, and Ben Mildenhall. 2022. DreamFusion: Text-to-3D using 2D Diffusion. *arXiv (2022)*.
- [6] Zhengyi Wang, Cheng Lu, Yikai Wang, Fan Bao, Chongxuan LI, Hang Su, and Jun Zhu. 2023. ProlificDreamer: High-Fidelity and Diverse Text-to-3D Generation with Variational Score Distillation. In *Advances in Neural Information Processing Systems*, A. Oh, T. Neumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Eds.), Vol. 36. 8406–8441.



A beautifully carved wooden queen chess piece.



A blue poison-dart frog sitting on a water lily.



A metal sculpture of a lion head, highly detailed.



A portrait of Batman, head, HDR, photorealistic, 8K.



A tarantula, highly detailed.



A portrait of Deadpool, head, HDR, photorealistic, 8K.



A camel with a colorful saddle.



A beautiful rainbow fish.



A red panda, HDR, photorealistic, 8K.



A puffin standing on a rock.

Figure 2: Generated 3D models of the proposed DreamLCM method.





A cat with azure eyes, HDR, photorealistic, 8K.



A DSLR photo of a sitting panda holding coco cola, photorealistic.



A cute hamster.



A border collie, photorealistic, 8K.



A cute Koala on a tree, HDR, photorealistic, 8K.



A gleaming silver trophy, HDR, photorealistic, 8K.



A snail on a leaf.



A sloth on a tree.



A beautiful suit made out of moss, on a mannequin.  
Studio lighting, high quality, high resolution.

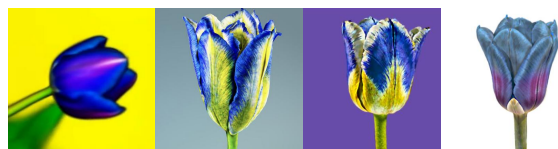


A baby dragon hatching out of a stone egg.

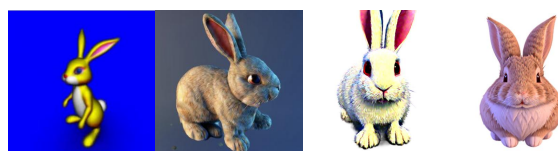
Figure 3: Generated 3D models of the proposed DreamLCM method.



A cauldron full of gold coins.



A high quality photo of a blue tulip.



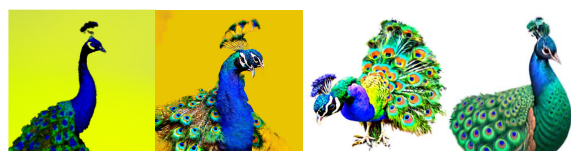
A rabbit, animated movie character, high detail 3d mode.



A rabbit, animated movie character, high detail 3d mode.



A chow chow puppy.



A peacock.



A roast turkey on a platter



A cute baby penguin.



A DSLR photo of a bald eagle.



A DSLR photo of a Schnauzer wearing a pirate hat.



A teddy bear sitting on a stone and wearing a scarf and wearing a flat cap.



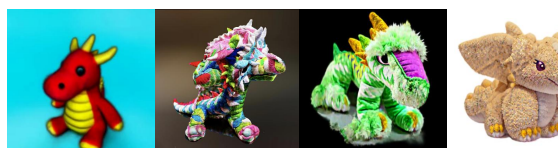
A pineapple.



A piglet in a teacup.



An overstuffed pastrami sandwich.



A plush dragon toy.



A human skull.

Figure 4: Comparisons with the current SoTA methods given the same text prompts.