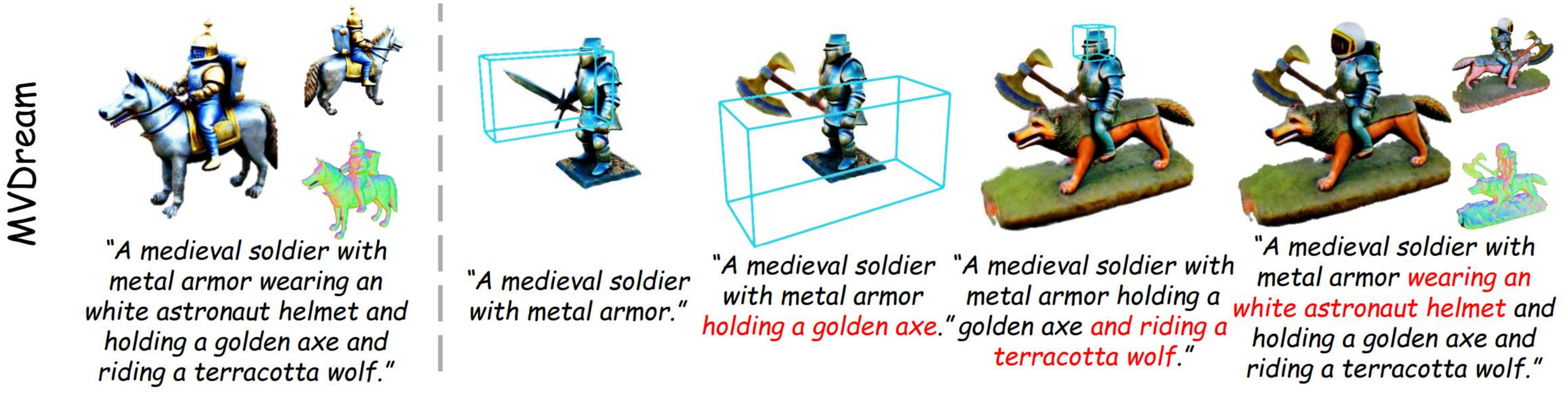
Progressive3D: Progressively Local Editing for Text-to-3D Content Creation with Complex Semantic Prompts

Xinhua Cheng, Tianyu Yang, Jianan Wang, Yu Li, Lei Zhang, Jian Zhang, Li Yuan
Peking University, International Digital Economy Academy (IDEA), Peng Cheng Laboratory

Challenge & Conception

(a) Object Missing (b) Attribute Mismatching (c) Quality Reduction

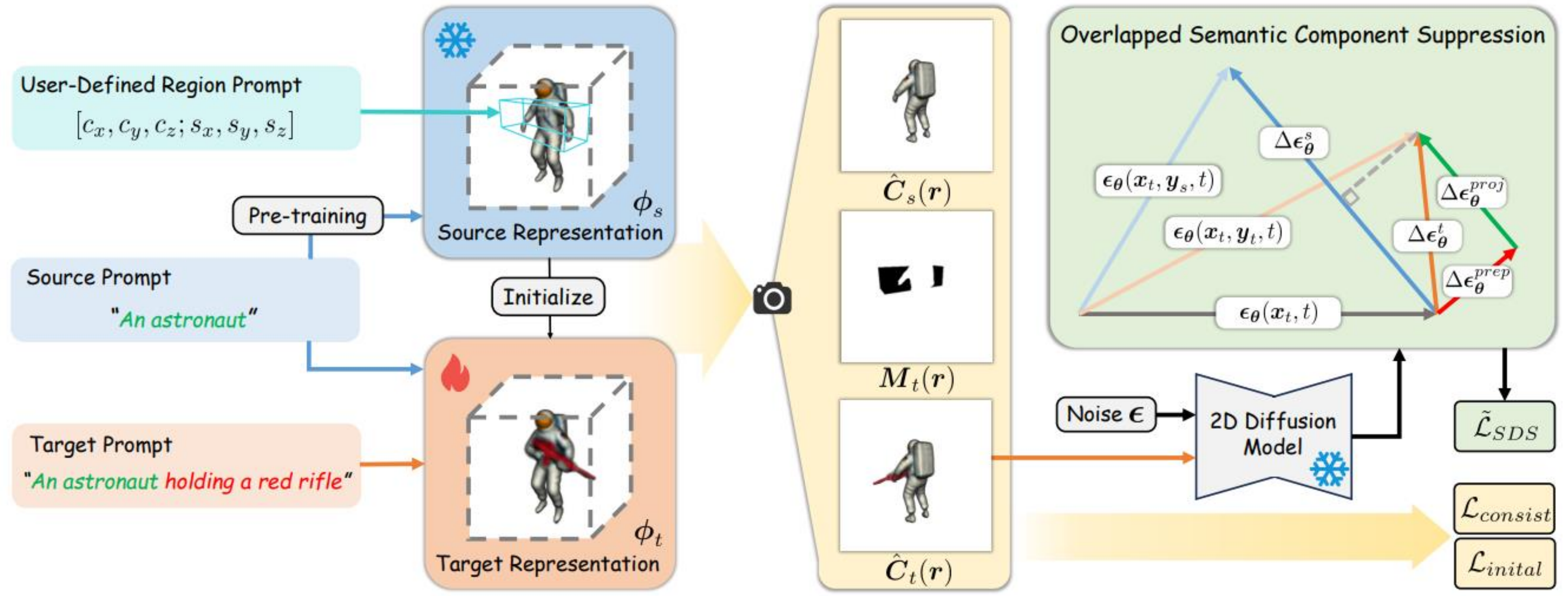
Challenge. Current text-to-3D methods often fail to produce precise results when the given prompt describes multiple interacted objects binding with different attributes, leading to significant issues including object missing, attribute mismatching, and quality reduction.



(a) Generating Directly (b) Generating with Progressive3D

Conception. We propose a framework named **Progressive3D** for creating precise 3D content prompted with complex semantics by decomposing a difficult generation process into a series of local editing steps. Compared to (a) generating with existing methods, (b) generating with Progressive3D produces 3D content consistent with given prompts.

Framework



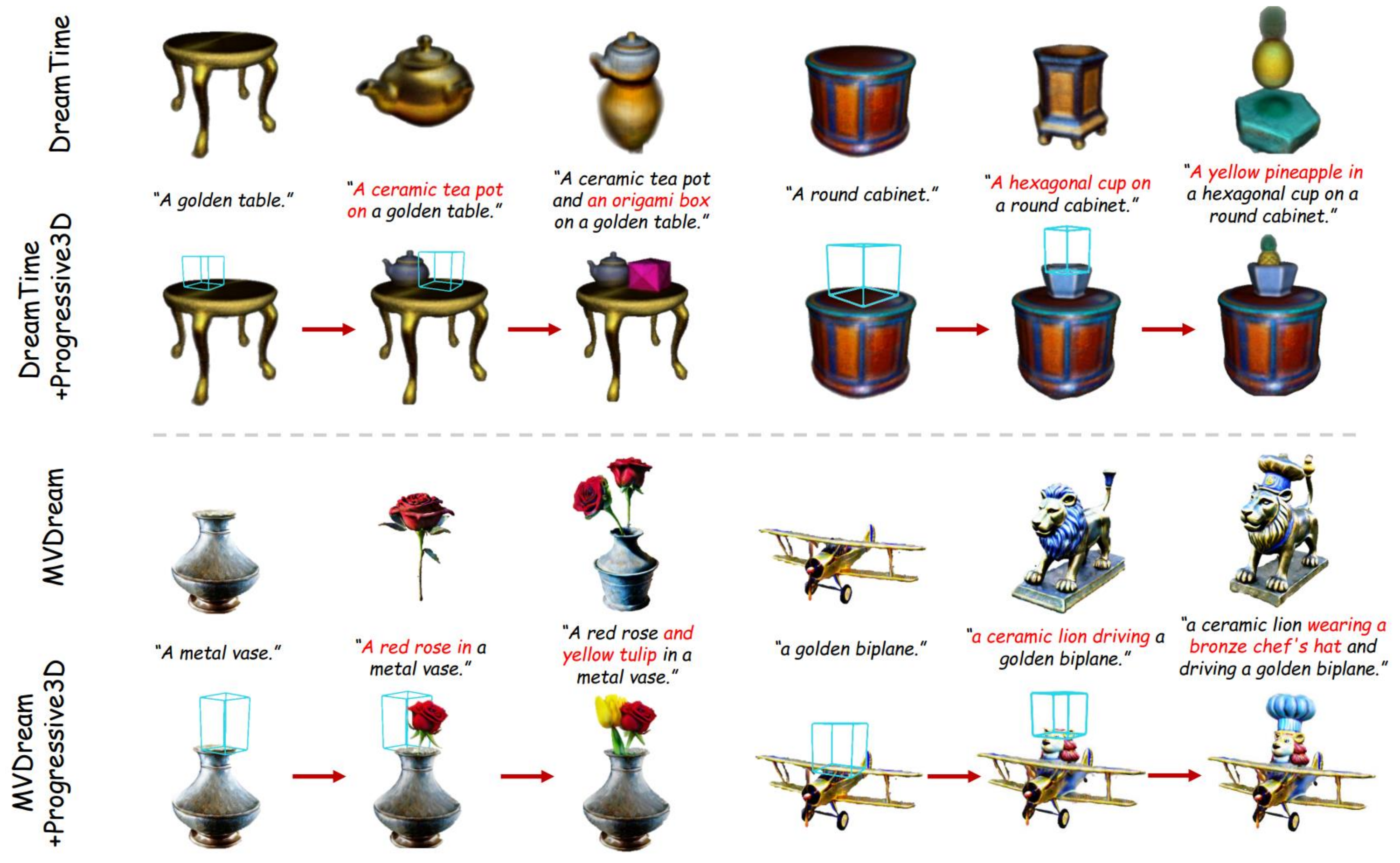
(a) Source Content and Region Prompt (b) 2D Mask M_t (c) w/o $\mathcal{L}_{consist}$ (d) w/o $\mathcal{L}_{initial}$ (e) w/o Overlapped Semantic Component Suppression (f) Ours

Consistency Constraint. Maintaining 3D content beyond user-defined editable regions unchanged.

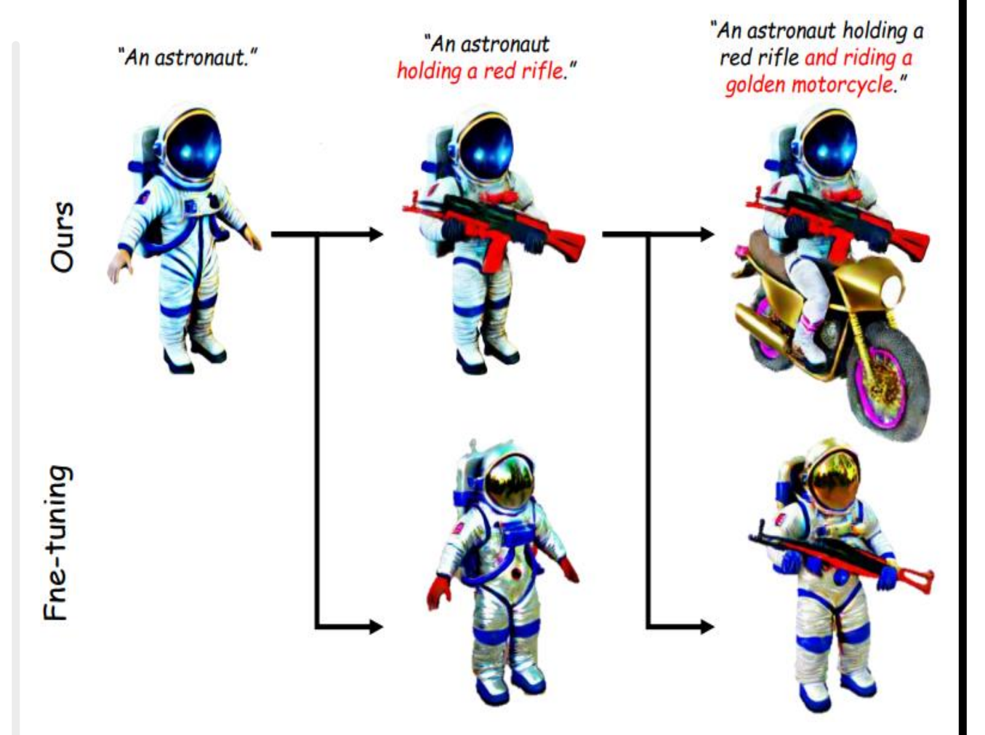
Initialization Constraint Encourage 3D content generated in empty space.

Overlapped Semantic Component Suppression. Discover the overlapped component between prompts, and enhance the influence of the different semantic by suppression.

More Experiments



Method	Metrics		Human
	B-VQA ↑	mGPT-CoT ↑	Preference ↑
DreamTime	0.227	0.522	16.8%
+ CEBM	0.186	0.491	-
+ A&E	0.243	0.528	-
+ Progressive3D	**0.474**	**0.609**	**83.2%**



Wechat **Code**