

LazyDrag: Enabling Stable Drag-Based Editing on Multi-Modal Diffusion Transformers via Explicit Correspondence

Zixin Yin, Xili Dai, Duomin Wang, Xianfang Zeng,
Lionel M. Ni, Gang Yu, Heung-Yeung Shum

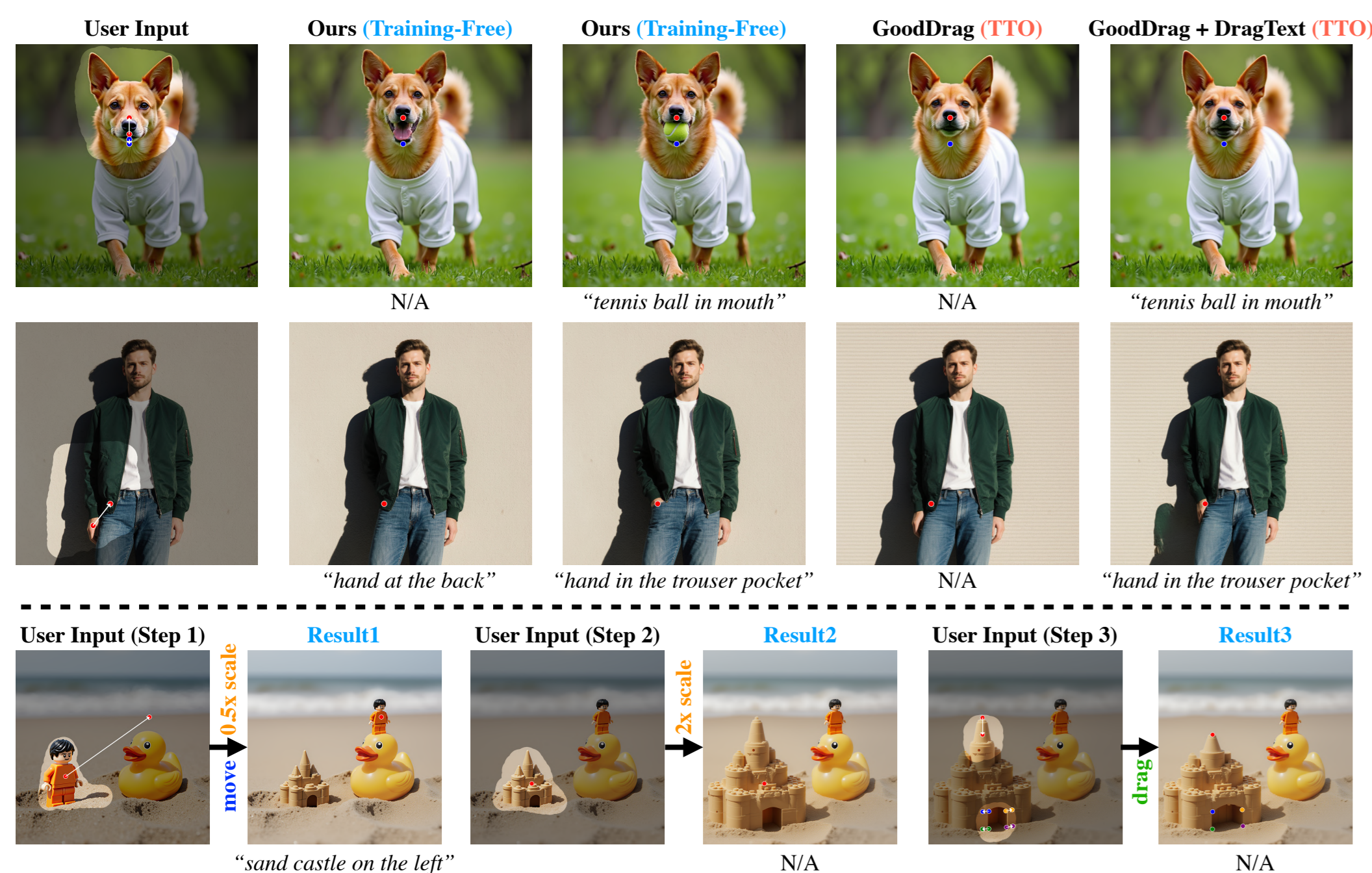


PDF

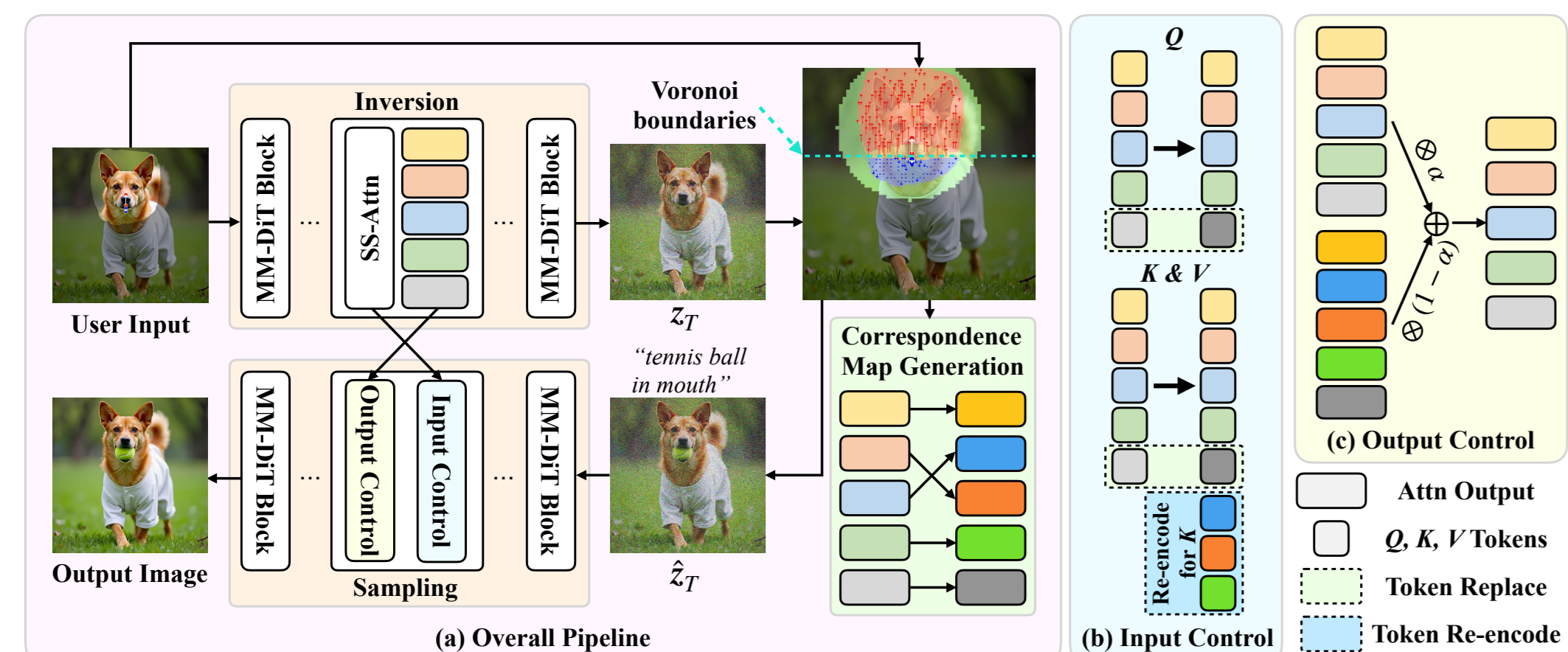


Website

Drag-Based Editing



Pipeline of ColorCtrl



Benchmark Result

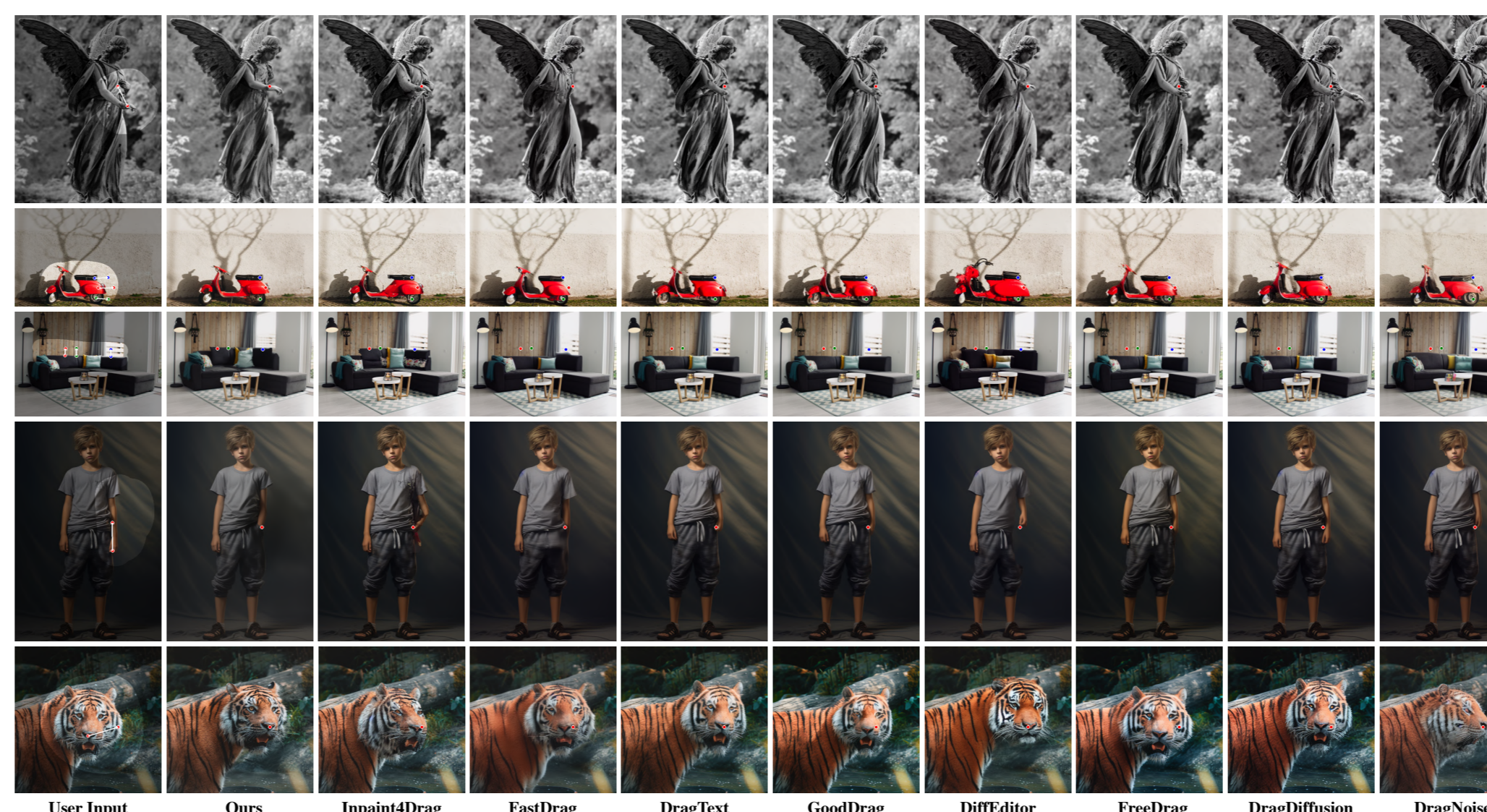
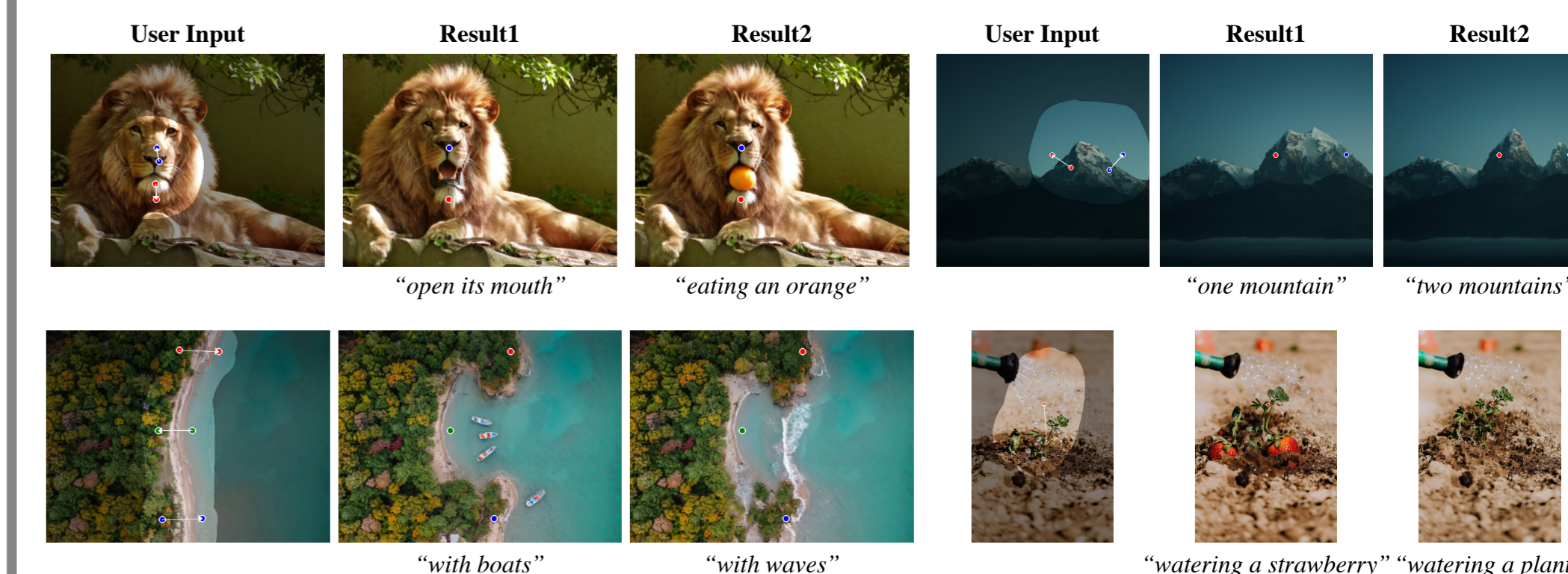


Table 1: Quantitative results compared with baselines on Drag-Bench.

Method	TTO-Req	MD ↓	SC ↑	PQ ↑	O ↑
DragNoise (Liu et al., 2024a)	✓	37.87 ± 0.23	7.793 ± 0.04	8.058 ± 0.01	7.704 ± 0.01
DragDiffusion (Shi et al., 2024)	✓	34.84 ± 0.30	7.905 ± 0.01	8.325 ± 0.02	7.798 ± 0.03
FreeDrag (Ling et al., 2024)	✓	34.09 ± 0.60	7.928 ± 0.02	8.281 ± 0.03	7.816 ± 0.02
DiffEditor (Mou et al., 2024a)	✓	26.95 ± 0.24	7.603 ± 0.01	8.266 ± 0.01	7.715 ± 0.01
GoodDrag (Zhang et al., 2025)	✓	22.17 ± 0.16	7.834 ± 0.03	8.318 ± 0.01	7.795 ± 0.01
DragText (Choi et al., 2025)	✓	21.51 ± 0.21	7.992 ± 0.02	8.227 ± 0.03	7.886 ± 0.01
FastDrag (Zhao et al., 2024)	✗	31.84 ± 0.96	7.935 ± 0.09	8.278 ± 0.01	7.904 ± 0.06
Inpaint4Drag (Lu & Han, 2025)	✗	23.68 ± 0.05	7.802 ± 0.06	7.961 ± 0.04	7.615 ± 0.06
Ours	✗	21.49 ± 0.04	8.205 ± 0.03	8.395 ± 0.03	8.210 ± 0.03

Cases with Various Additional Text Prompts



Drag and Move Mode

