

Objaverse++: Curated 3D Object Dataset with Quality Annotations

Supplementary Material

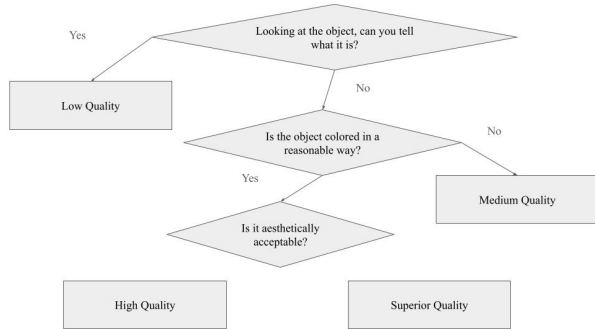


Figure 9. Decision tree for human annotators to categorize the quality level of a 3D object.

7. Quality Score Rubrics

Due to the specialty of judging 3D objects, we composed additional training material for our human annotators, including Figure 9 and the rubrics below. For quality score, here are some criteria to consider when an object has proper semantic meaning and texture.

High-Quality Criteria:

- Basic color scheme present but lacks richness and aesthetic appeal.
- Acceptable geometric shapes - not too rough but not highly detailed.
- Basic textures present - goes beyond flat surfaces but lacks sophistication.
- Visually comfortable and harmonious, but lacks refinement in details (like color rendering and fabric textures).

Superior-Quality Criteria:

- High-quality modeling with rich textures, vibrant colors, and aesthetic value.
- Rich, harmonious color combinations that feel natural or appropriate to the style.
- Geometric proportions that match either real-world references or suit the intended artistic style.
- Detailed surface texturing with effective lighting/shading.
- Aesthetically pleasing or visually impactful.
- Abundant detailed elements such as decorations, patterns, etc.

8. Training Loss

Figure 10 demonstrates the impact of dataset quality on training loss. The high and superior quality subsets show faster and more stable convergence than the baseline, a randomly sampled subset of 100,000 objects from Objaverse, and random 50k subsets. Quality-filtered data re-

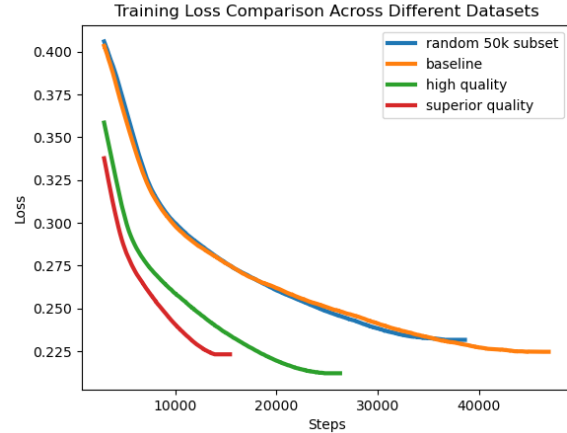


Figure 10. Training loss comparison across different datasets. Similar to the validation loss result 8, the model converges significantly faster on high-quality and superior-quality datasets, and converges roughly at the same speed on a random 50k subset and baseline, which is a randomly sampled subset of 100,000 objects from Objaverse.

	Our model	Baseline	No preference
Building	31.9	48.9	19.2
Ceramic	51.1	38.3	10.6
Fire	46.8	40.4	12.8
Girl	95.7	0	4.3
Hotdogs	27.7	12.8	59.5
Hydrant	61.7	31.9	6.4
Lamp	61.7	29.8	8.5
Mailbox	44.7	48.9	6.4
Owl	51.1	19.1	29.8
Traffic	46.8	12.8	40.4

Table 4. User study results in percentage. Of the 10 pairs of objects, 8 preferred our model (in bold) over the baseline.

duces noise, accelerates optimization, and enhances learning stability, allowing the model to converge more efficiently. In contrast, the baseline dataset’s noisy samples hinder optimization. The superior quality subset achieves the best results among the four datasets, underscoring the importance of high-quality data over dataset size for efficient model training.

9. User Study Results

We include an additional Table 4 for the percentage breakdown of our user study.