

Appendix

Screenshot of Human evaluation website

Caption

First Image/ Left Image

An image shows an apple-shaped terrarium to the right of a pear-shaped terrarium.



Does the First Image satisfy the caption?

YES NO

Generated Images/ Right Images



Naturalness: Is this image Natural? YES NO

Relevance: Is the image Relevant to the caption? YES NO

Correctness: Is the image Correct? YES NO

SUBMIT