

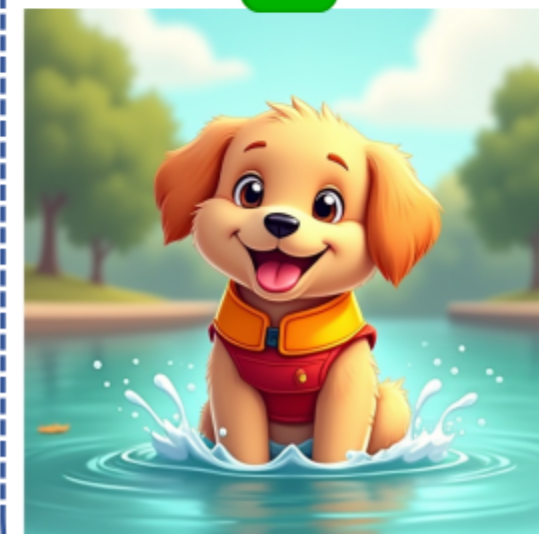
ConsistencyRank

Rank the images
by **Consistency**



Reference

Pair Images



1



3

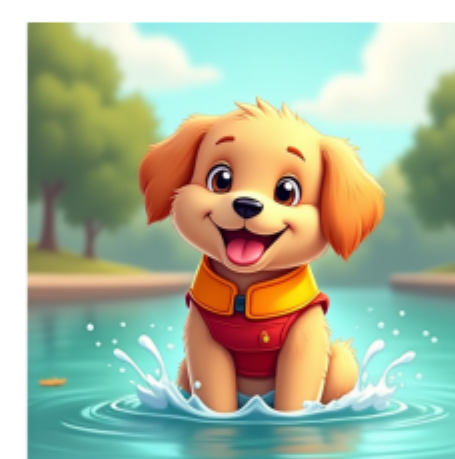


2



4

Given two images under *the same theme* {prompt},
are they consistency in all aspects such as style,
logic and ideneity? Answer “Yes” and “No” first,
then provide 1~3 detailed reasons.



ConsistencyReward-Dataset

Yes. 1. Both images depict the **same** cartoon-style golden **puppy** with **identical** facial features, furcolor, and expression. 2. The **artistic style**—soft lighting, warm colors, and smooth shading—is consistent across both scenes.

No. The two puppies differ in **identity**. Though both depict similar golden puppies, their facial features, ear shapes, and fur tones differ, suggesting they are not the same dog.

Wow! It learns preference
through binary labels



Image 1 Tokens



Image 2 Tokens



Prompt Tokens

ConsistencyReward-7B

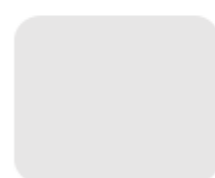


Training

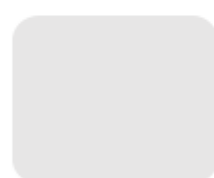
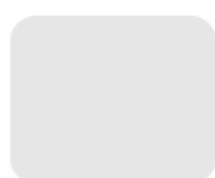
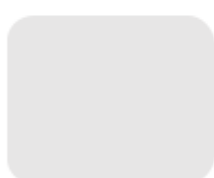
Evaluation



Yes/No



Output Reason Tokens



Cross Entropy Loss



Yes/No



Label Reason Tokens

