Rejected Image Preference

Rejected Instruction Preference

Rejected Response Preference

Where are the man and **the lion** located in this image?

The man and **the lion** are located on the beach, near the ocean.

Where are the man and **his dog** located in this image?

The man and **his dog** are located on the beach, near the ocean.

Chosen Preference

Collected Multimodal Preference

Entity-centric Multimodal Preference Optimization (EMPO)

Final LVLM

(1) Multimodal Preference Construction

(2) Multimodal Preference Alignment