

Modality 1: Text

This is a red apple.

A coffee mug.

Modality 2: RGB



Modality 3: Depth

