

Q Multimodal LLMs

Solution

process and generate
content across modalities

Eg. text to image

Visual Question Answering