

SOTA Models for Instruction Following & Visual Dialogue



Figure 12: Multimodal Chat/Agents