















Text-to-Video generation: "a horse galloping on a street"

















Text-to-Video generation: "a panda is playing guitar on times square"















Text-to-Video generation + pose control: "a bear dancing on the concrete"















Video Instruct-Pix2Pix: "make it Van Gogh Starry Night style"

Text-to-Video generation + edge control: "white butterfly"