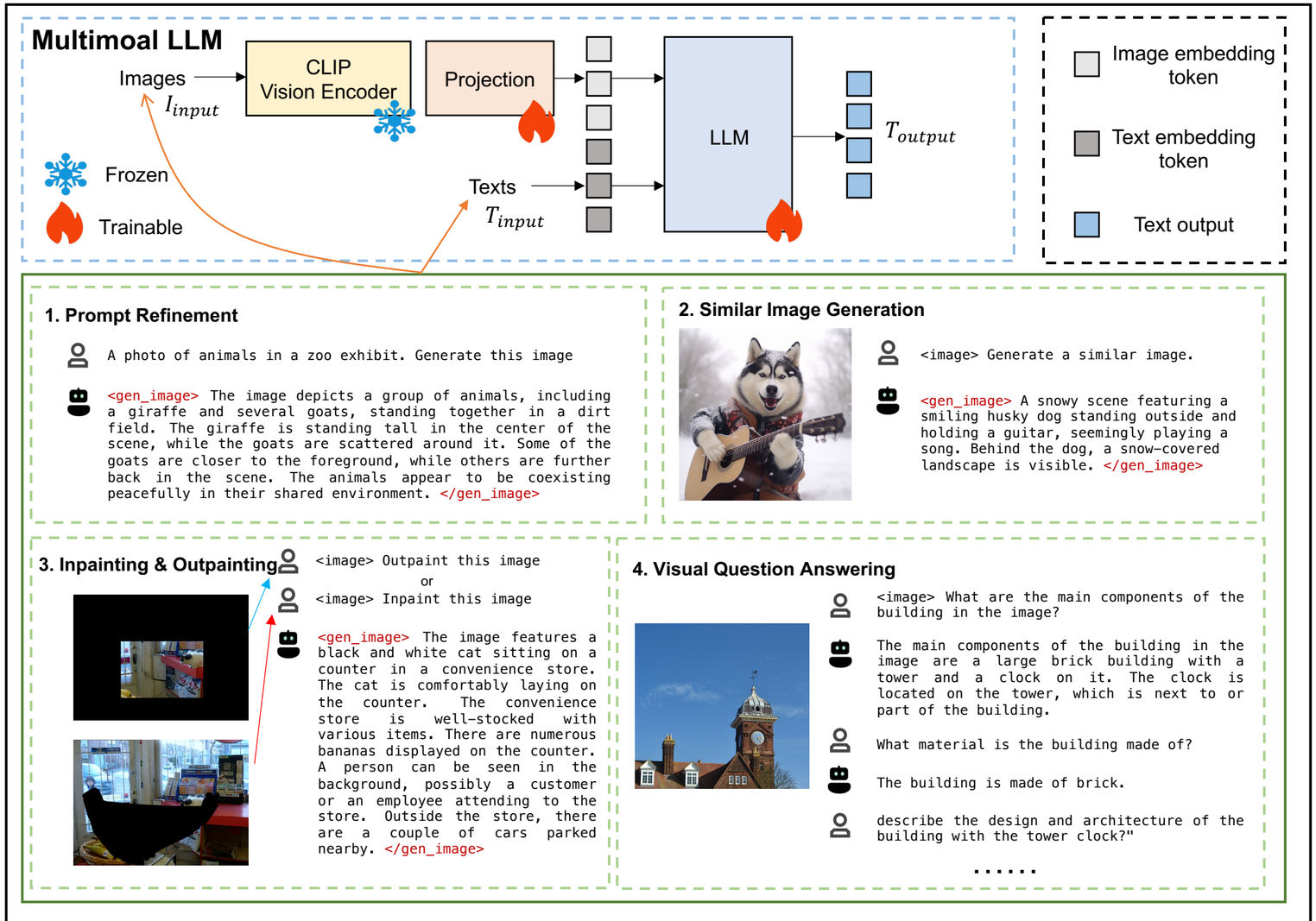
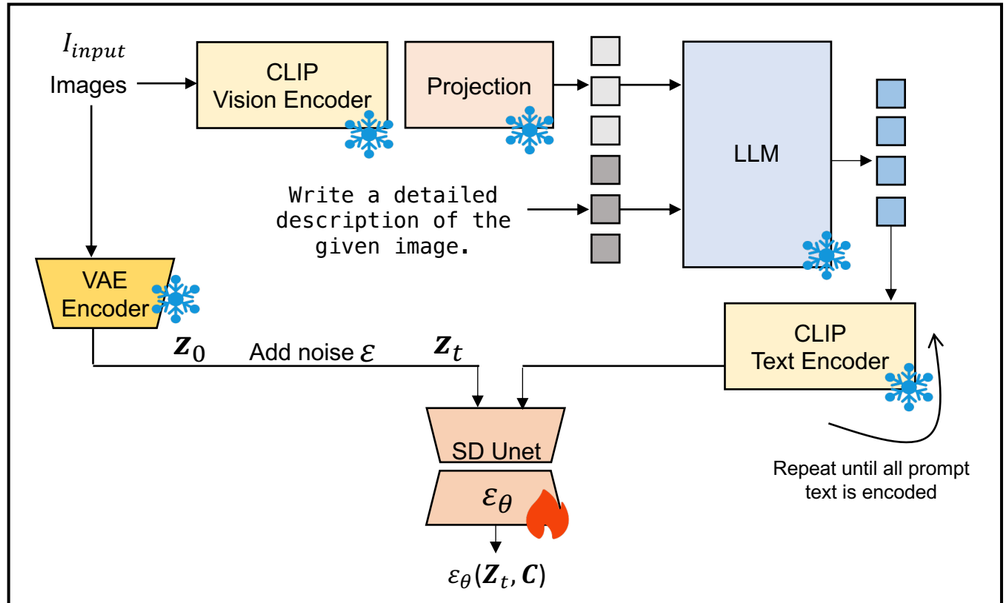
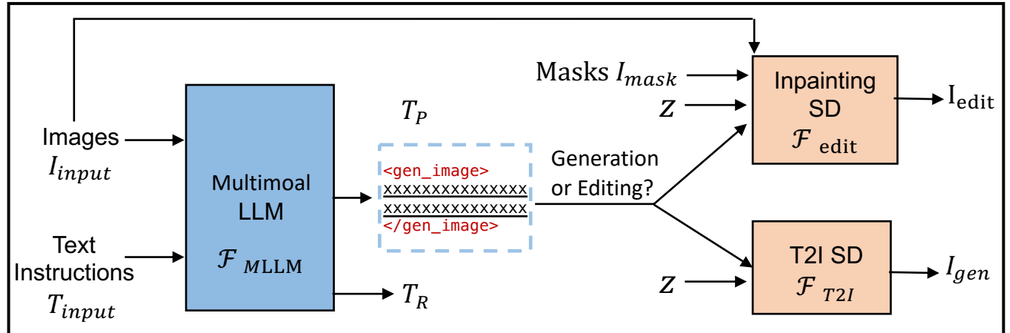


(a) Stage 1: Training multimoal LLM for image generation



(b) Stage 2: Training Stable Diffusion adapted to LLM instructions



(c) Inference process of LLMGA