

Generation

SFT images

Prompts

SAM finetuned

SAM Mask
Generation

Per-prompt Responses

Annotation

SFT images

Human
Ranking

Ranked Responses

Optimization

SFT Policy

DPO

Optimized
policy

