

Multiple Concept Personalization

Team: CHILLGUY 電子所一 R13943013林沛翰 R13943018吳庭毅 R13943021陳品睿 電機四b10901073黃挺維

Brief Introduction

- Adjusted from Mix of Show
- LoRA based Stable Diffusion model base v1.5 single concept finetuning
- Gradient Fusion for different concept weights
- Region-control Sampling

SAM mask generation

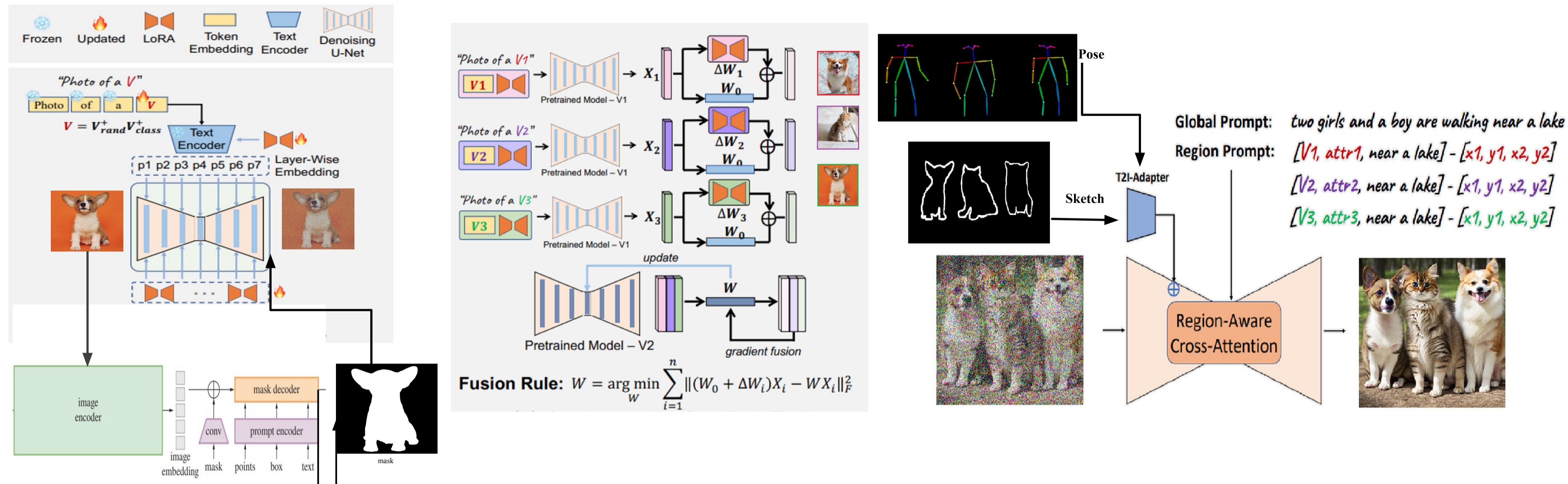
- we use SAM to quickly generate concept mask for cross attention

Ablation study about sketch control

Ablation study about pose control

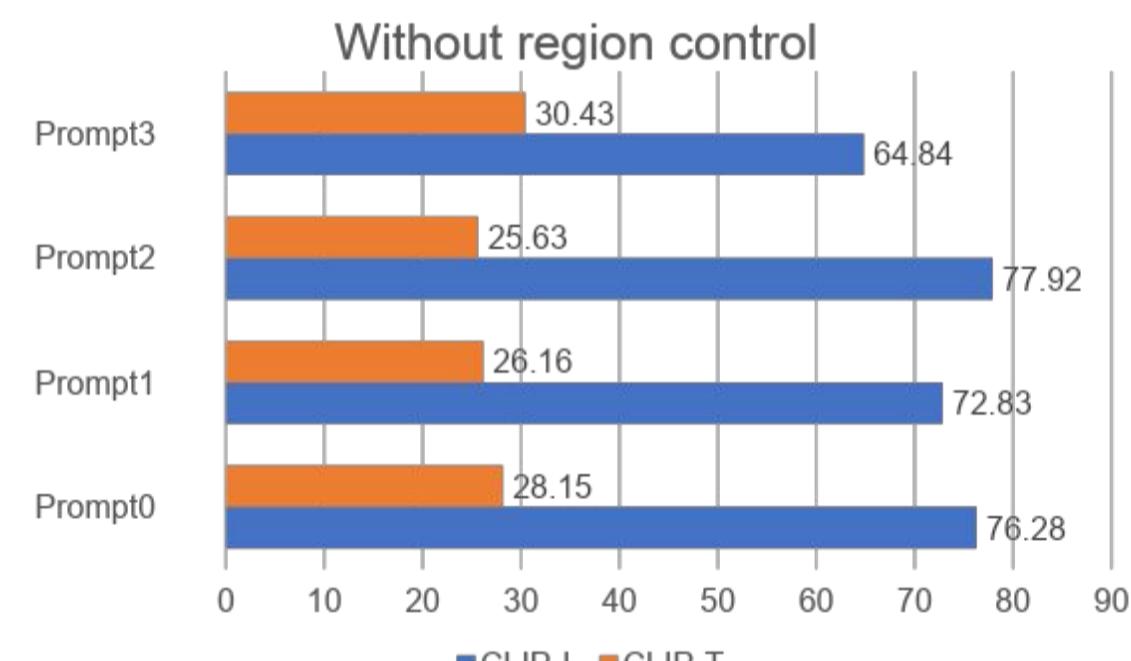
Other interesting experiments like OMG

Model Architecture

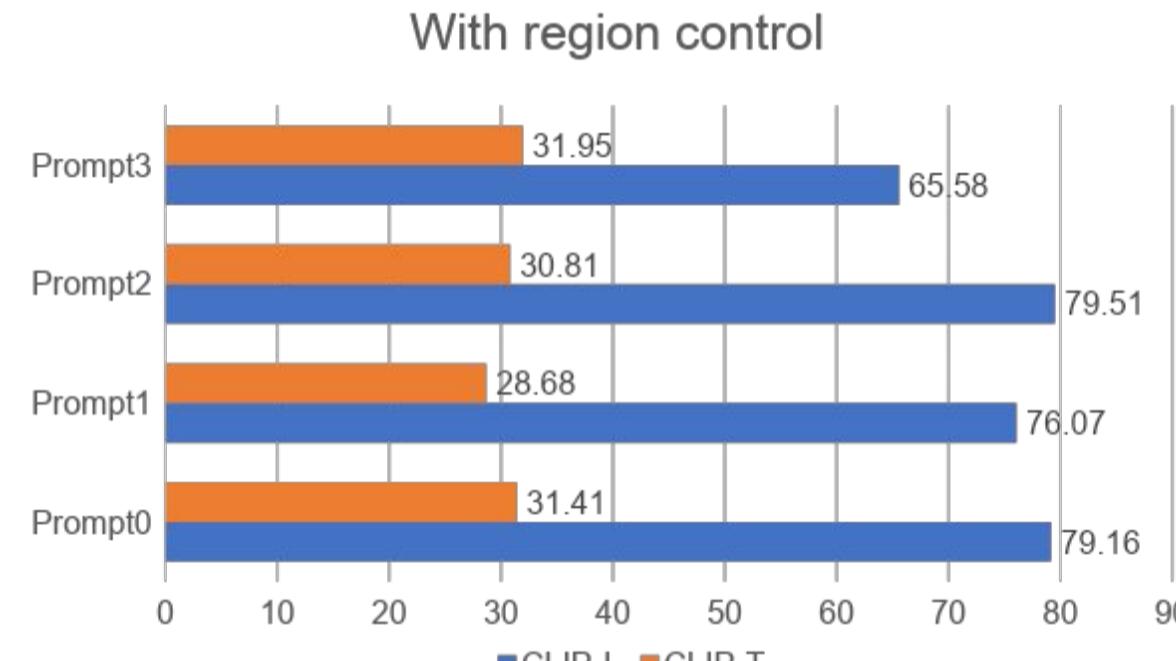
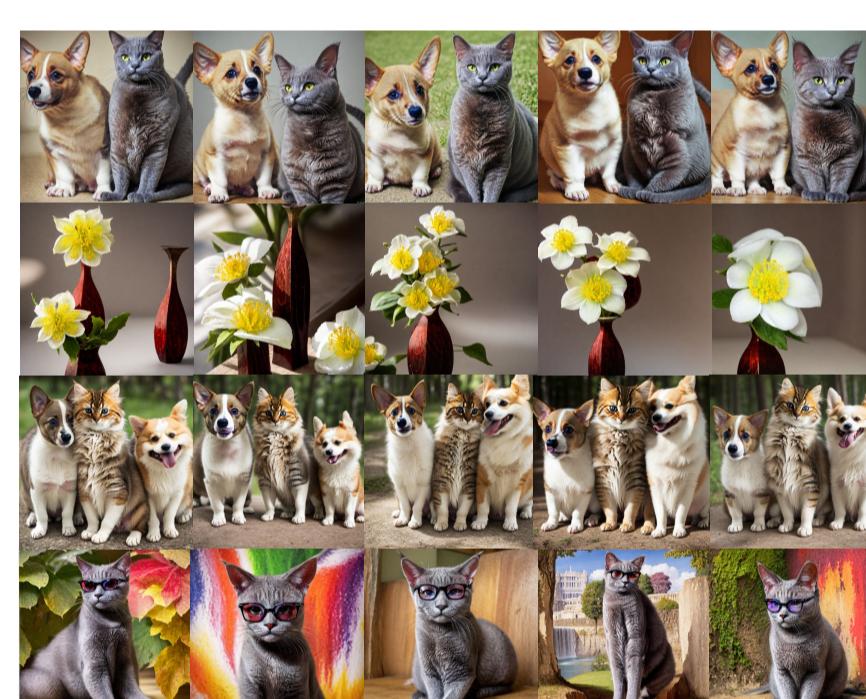


Ablation Study

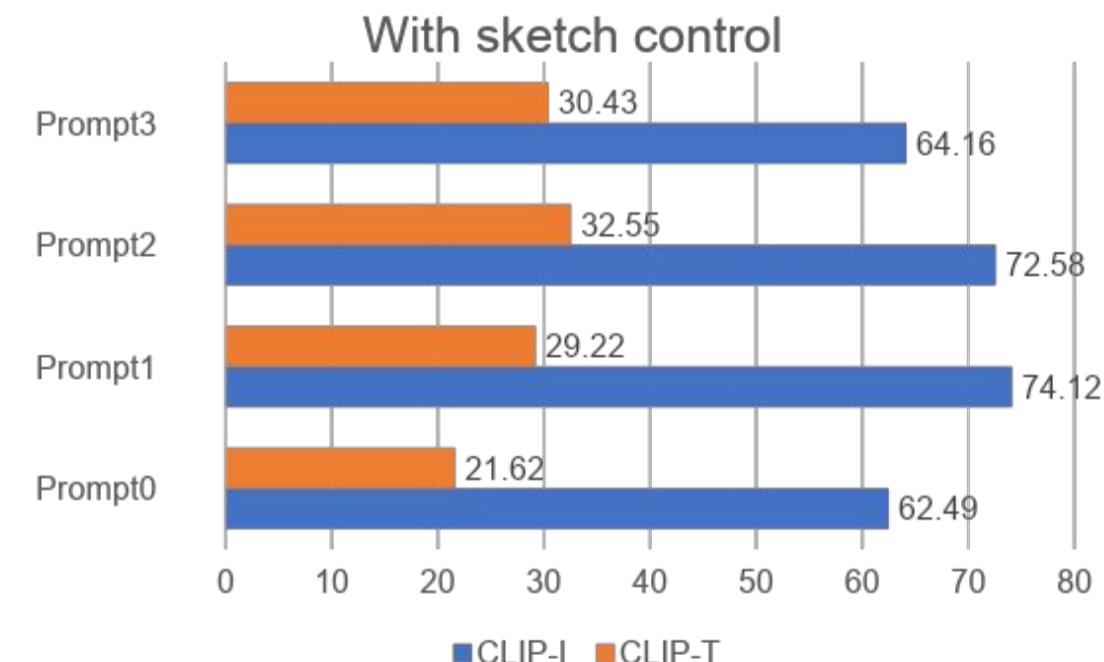
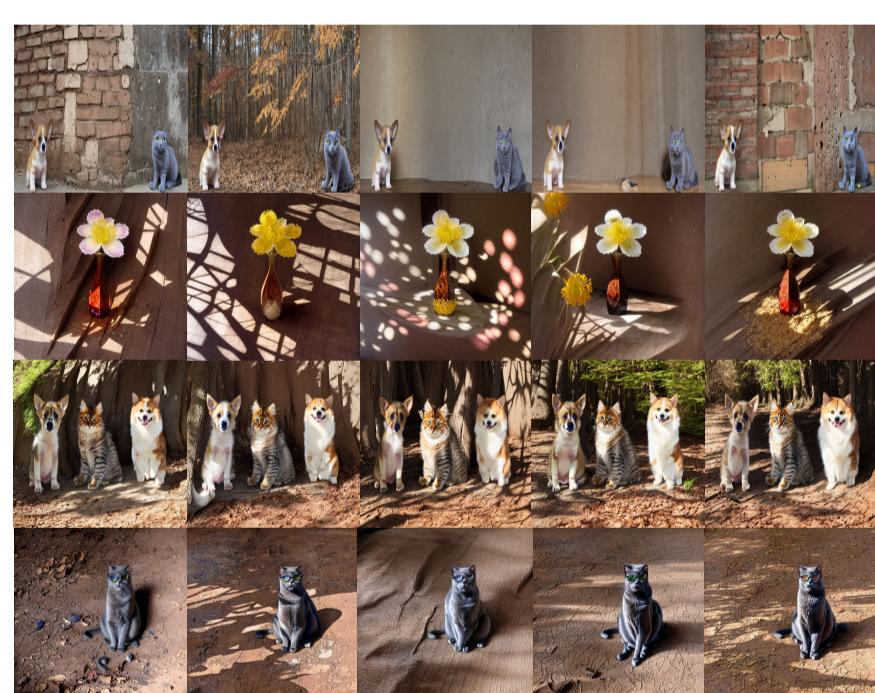
Without region control



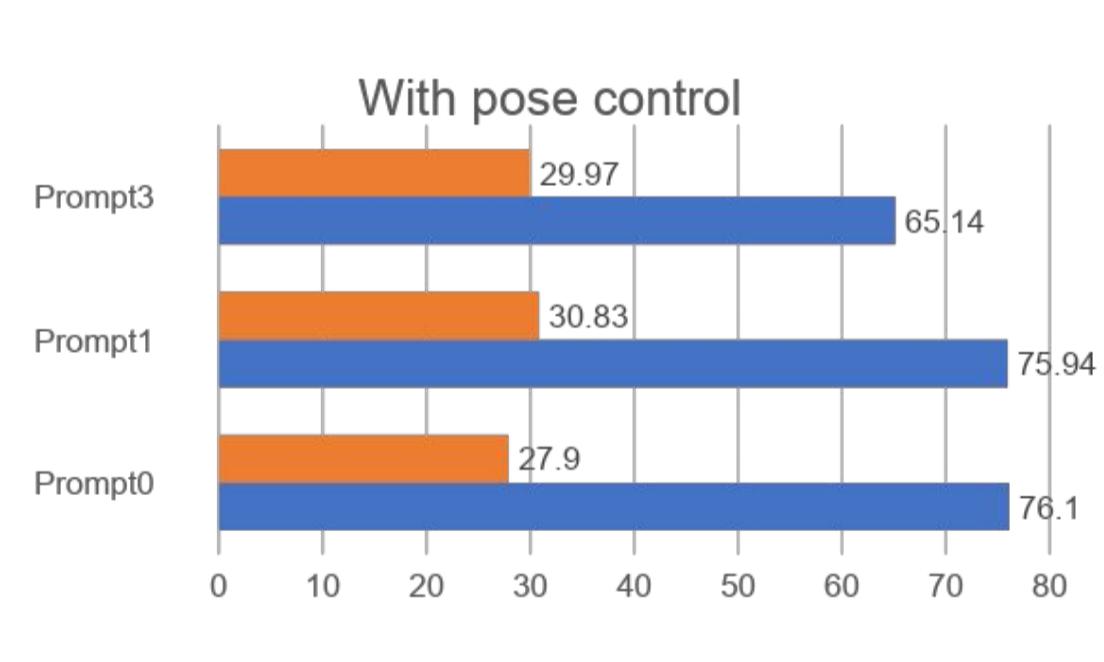
With region control



With sketch control



With pose control



Other Experiment - OMG

- two-stage sampling
- SDXL base 1.0 + LoRA + Canny ControlNet
- YOLO-World + EfficientViT SAM

Why

- no need for model merging or additional tuning (just lora)
- illumination harmony

But

- low success rate for non-human subjects (shape not match)

Solution & Trade off

- canny controlnet → loss of variety

