

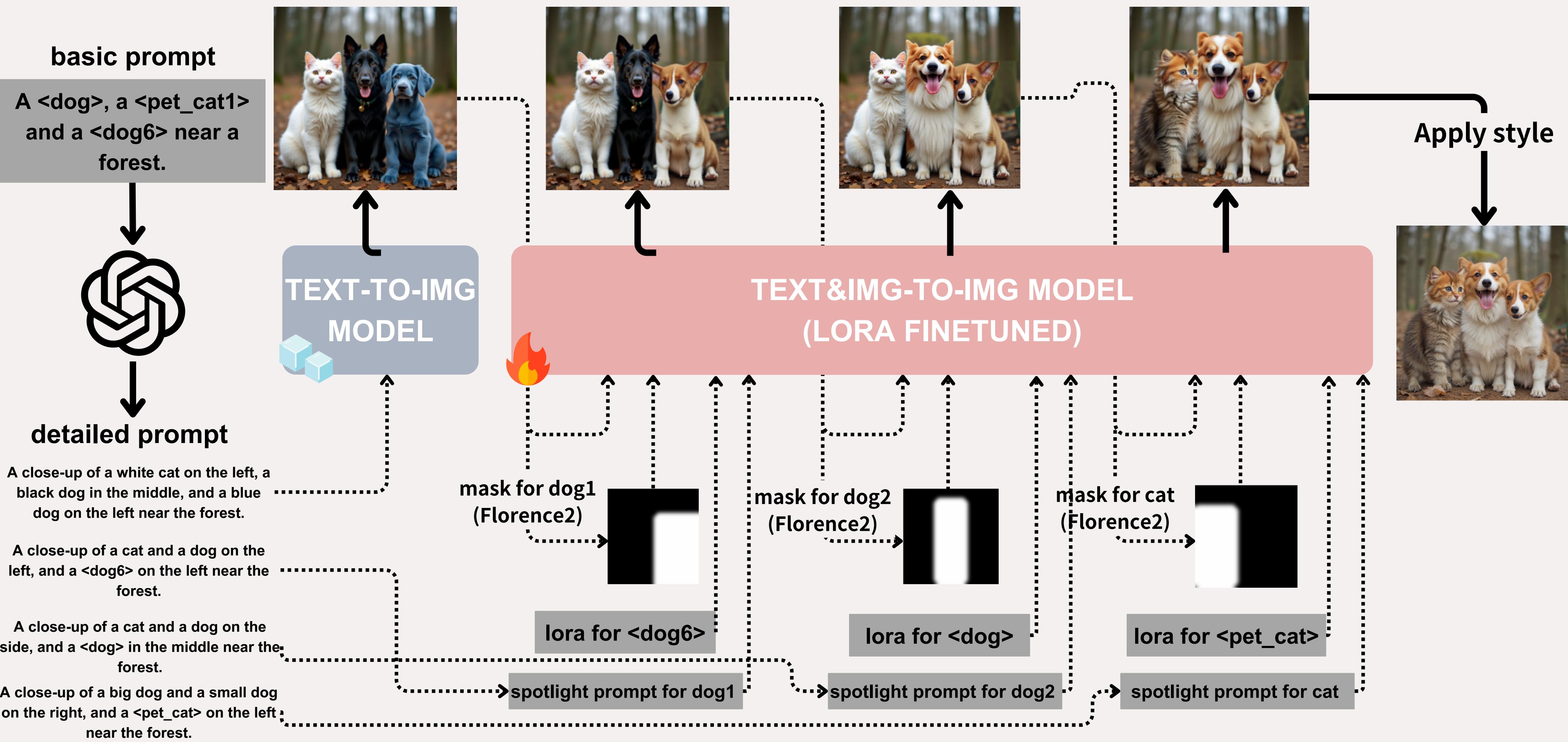
StageLight: Sequentially Targeted and Guided Editing for Layered Inpainting and Generative Hierarchical Transitions

B10901184 錡亭勳 B10901058 傅渝翔 B10901163 張顥譽 B10901176 蔡弘祥

ABSTRACT

Decomposing the prompt for multi-layer inpainting using LLM can significantly improve the model's capability, including creativity, precision, and customization similarity.

ARCHITECTURE



EXPERIMENT

StageLight(Ours)	OmniGen	Custom-Diffusion	Textual Inversion	Reference	Baselines
				A <cat2> on the right and a <dog6> on the left. <cat2> <dog6>	<ul style="list-style-type: none">Textual Inversion<ul style="list-style-type: none">Find a special token to represent the reference images.Designed for single concept.Custom Diffusion<ul style="list-style-type: none">Find a special token to represent the reference images.In addition, fine-tune the cross-attention layer of the stable diffusion.Designed for multiple concepts.OmniGen<ul style="list-style-type: none">Zero-shot generates images with special concepts.Designed for multiple concepts.
				A <cat2> wearing <wearable_glasses> in a <watercolor> style. <wearable_glasses> <watercolor>	
CLIP-I : 75.20 CLIP-T : 33.18	CLIP-I : 74.77 CLIP-T : 31.97	CLIP-I : 71.51 CLIP-T : 29.72	CLIP-I : 70.80 CLIP-T : 29.53		
Interaction	Interaction	Interaction	Camera	Camera / Depth	Style
Features	<ul style="list-style-type: none">InteractionCameraDepthStyle				

ABLATION STUDY

LD+SP(Ours)	W/o LD	SP on <dog>	W/o SP on <dog>	LD : Layer Decomposition SP : Spotlight Prompt	Reference
Prompt: A <dog>, a <pet_cat1> and a <dog6> near a forest.					