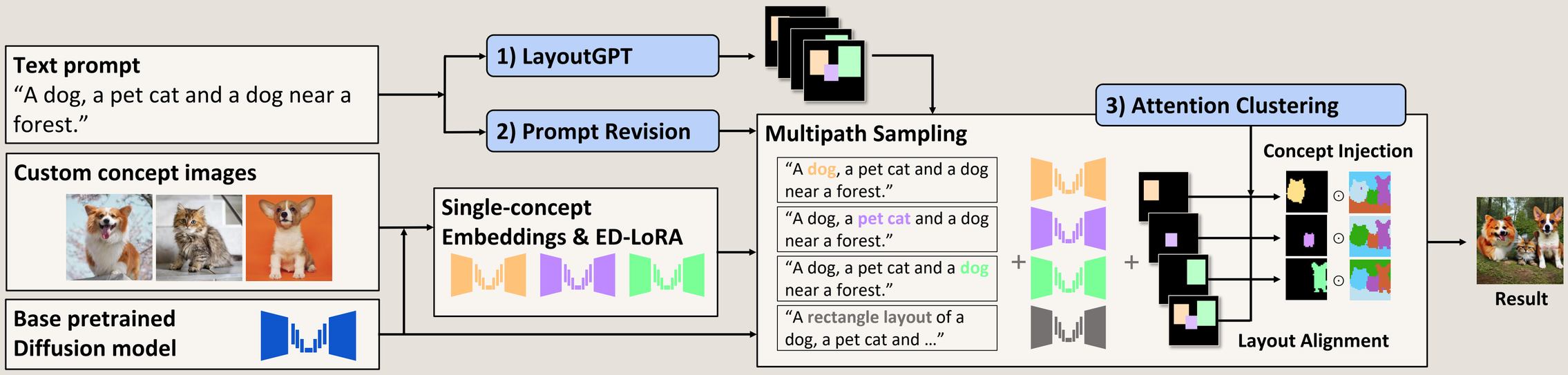
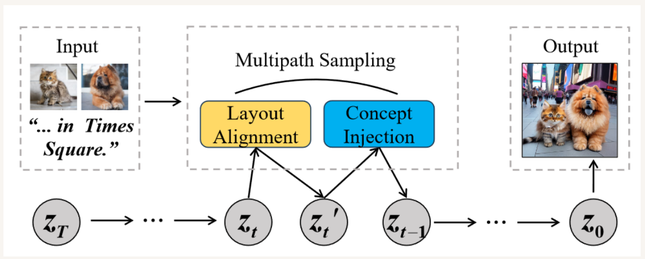
FRANC: Fusion of RANdom Concepts

Group 7: DontLookCanVis
莊承霖 (電子所二 R12943017) 葉惟欣 (資訊所一 R13922043) 韓家典 (ICDA一 R13k41014) 鄭宇彤 (醫學五 B09401064)



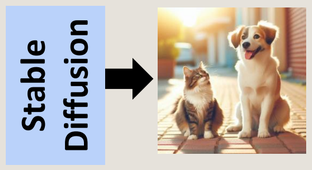
Introduction



Model: Concept Conductor

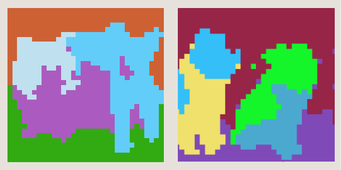
- **Layout Alignment**: Align with a **reference image** to reduce concept mixing.
- **Concept injection**: Inject attention features of single custom concept based on individual mask.

Motivation 1



Extra diffusion for reference image

Motivation 2

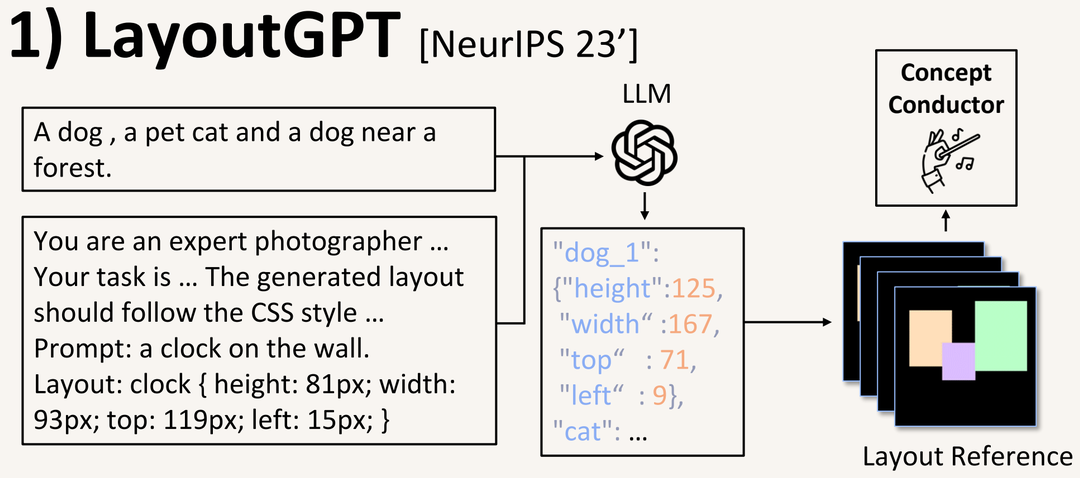


Incorrect masks during concept injection

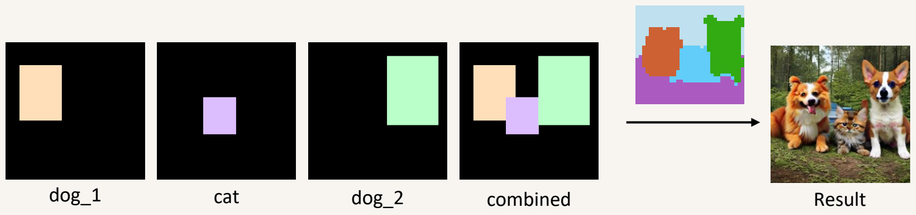
Motivation 3

Object replacement due to custom style concept

1) LayoutGPT [NeurIPS 23’]

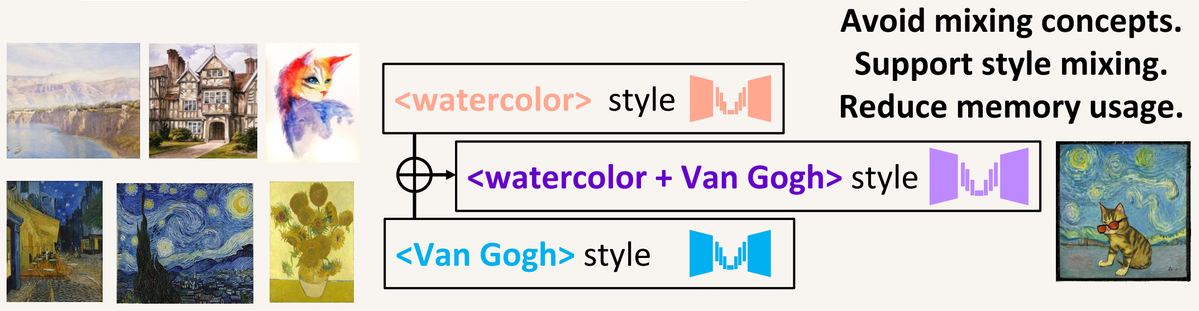


Our prompts consist of **task instructions**, and in-context examples in **CSS structures** to enhances LLMs’ interpretation of the spatial meaning and improve generation accuracy.

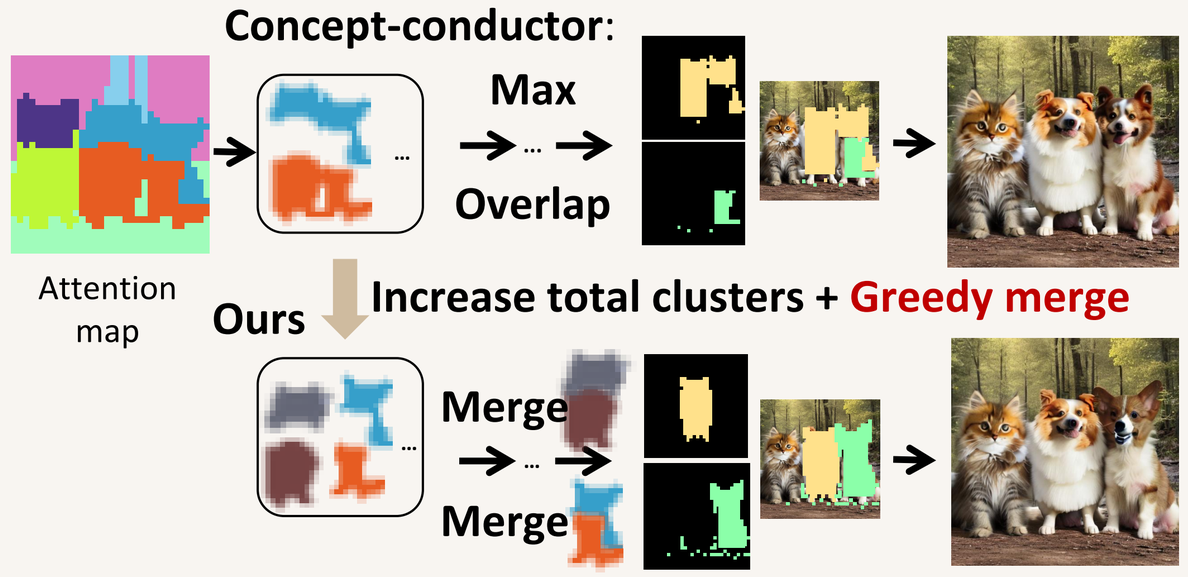


2) Prompt Revision

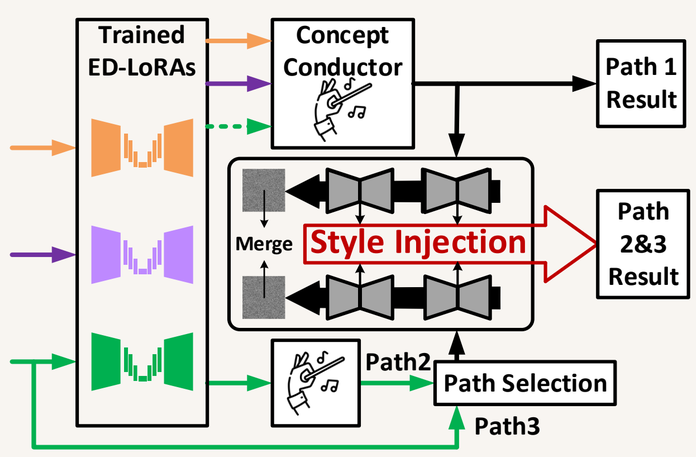
Avoid mixing concepts.
Support style mixing.
Reduce memory usage.



3) Attention Clustering



4) Ablation: Style Injection [CVPR 24’]



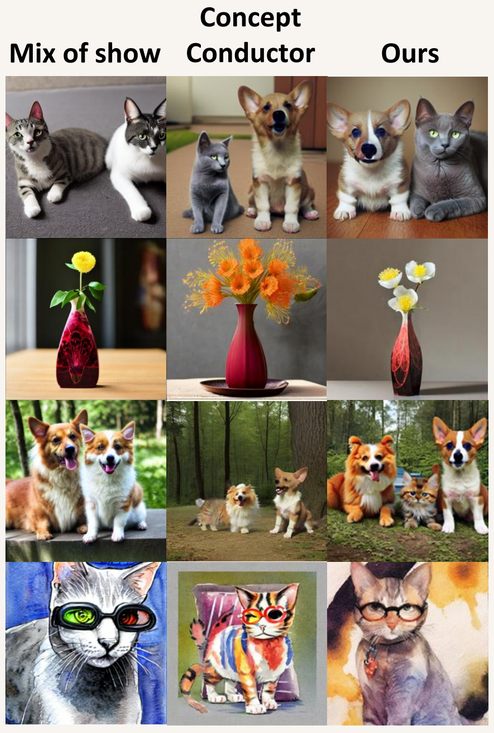
Diffusion-based style injection invert two images back to noises, merge them and **replace Q** with **content features**, **KV** with **style features** on each step.

Path 1 (without Injection)

Path 2 (Injection w/ trained images)

Path 3 (Injection w/ style images)

Result and Conclusion



#	2	3	4	CLIP-I	CLIP-T
Ours	+	-	-	74.30	33.20
	+	+	-	74.24	33.08
	+	+	+	74.09	33.11
Concept Conductor				72.47	31.70
Mix of show				72.73	30.29

- **LayoutGPT** remove the limitations of reference images and masks.
- **Prompts Revision** avoid mixing concepts.
- **Attention Clustering** resolves shared characteristics among different concepts in self-attention clusters.
- Integrate **style injection** to observe the model's handling of the style concept.

Reference

Concept Conductor
Mix-of-Show
LayoutGPT
StyleID