

Using Generative Models to Improve Generative Models

Yubei Chen
Assistant Professor
ECE Department, UC Davis
Cofounder, Aizip.ai



Personalizing Text-to-Image Diffusion Models



Input images



in the Acropolis



swimming



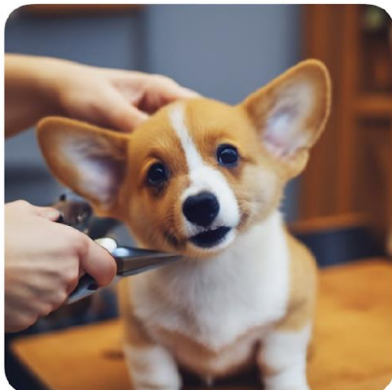
sleeping



in a doghouse



in a bucket



getting a haircut

(Dreambooth, CVPR'23)

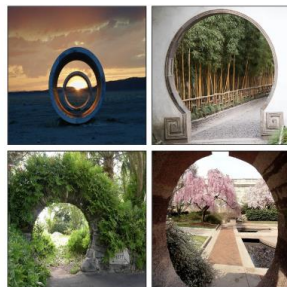
Personalizing Text-to-Image Diffusion Models



Input images



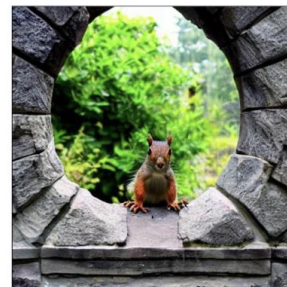
in the Acropolis



A photo of a **moongate**



A **moongate** in the snowy ice



A squirrel in front of **moongate**



swimming



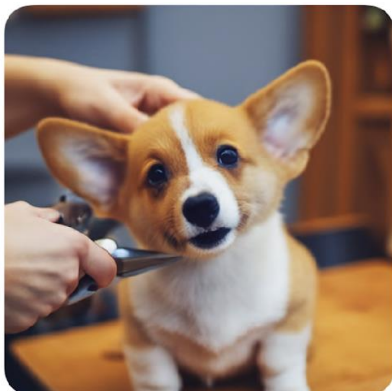
sleeping



in a doghouse



in a bucket

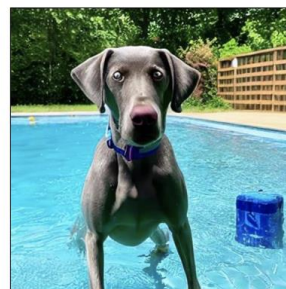


getting a haircut



A photo of a **V* dog**

User input images



A **V* dog** in a swimming pool



A **V* dog** wearing sunglasses

Single-concept generation

(Custom Diffusion, CVPR'23)

(Dreambooth, CVPR'23)

Failures of Current Models: Precise Instruction Following



Prompt = V_1^* dog and V_2^* cat playing together

Dreambooth



Custom Diffusion

Ours (joint training)



V_1^* dog and a V_2^* cat
playing together

Failures of Current Models: Precise Instruction Following



Prompt = V_1^* dog and V_2^* cat playing together

Dreambooth



Custom Diffusion

Ours (joint training)



V_1^* dog and a V_2^* cat
playing together

Pretrained model



dog and a cat
playing together

Stable Diffusion 1.5



Misalignment Between Text and Image in the Dataset



Corner Bakery Sweets Jigsaw Puzzle



George K Ralph Evening New York



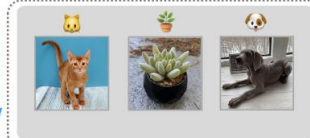
ON SALE Country Hutch in Greens,
Yellows and Reds for 1:12 Scale Dollhouse



Ben's Confectionary (Looking Back) (HOL770205),
a 1000 piece jigsaw puzzle by Holdson. [Click to view larger image.](#)

To Use Generative Models to Create High-Quality Data

1 Source Personal Images

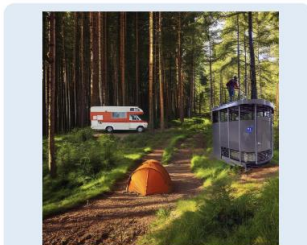


→ Gen4Gen → Text-to-Image Diffusion

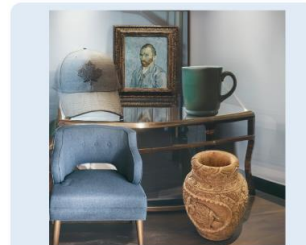
2 MyCanvas Dataset



caption: In the garden with colorful flowers, a 🐶, a 🐱, and a 🌿 next to each other.



caption: A 🚐 parked by 🚌 and 🏕️ in woods. The forest is lush with trees, and the sky is visible through the canopy.

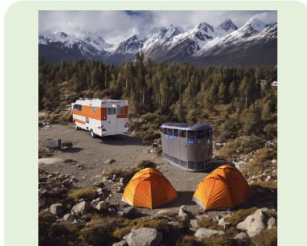


caption: A 🖼️ of a man is on a table next to a 🪑 and a 🍵. A 🪑 is also in the room. The 🏺 is on the floor.

3 Multi-Concept Generation



A night view of Times Square, a 🐶 and a 🐱 walking near few 🌿.



🚐 is next to a 🚌, and two 🏕️ is located with rocks, near snowy mountains.



A 17 century hall, a 🖼️ on wall. One 🪑 and one 🍵 on table, and one 🪑 and one 🏺 on the floor.

To Use Generative Models to Create High-Quality Data

User Images



To Use Generative Models to Create High-Quality Data

1

Object Association & Foreground Segmentation

Source Objects

Segmented Objects

$O' = \{\text{🐶} \text{🌿} \text{🐱}\}$

$\mathcal{M}(\mathcal{D}(\text{🐶}))$



$\mathcal{D}(\text{🐶})$

$\mathcal{M}(\mathcal{D}(\text{🌿}))$



$\mathcal{D}(\text{🌿})$

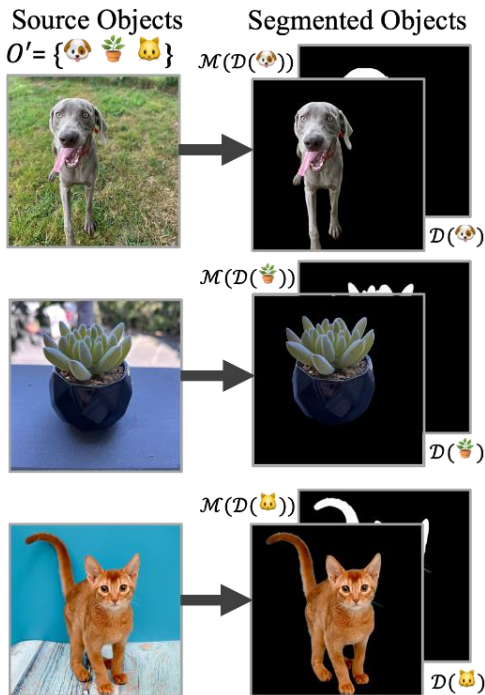
$\mathcal{M}(\mathcal{D}(\text{🐱}))$



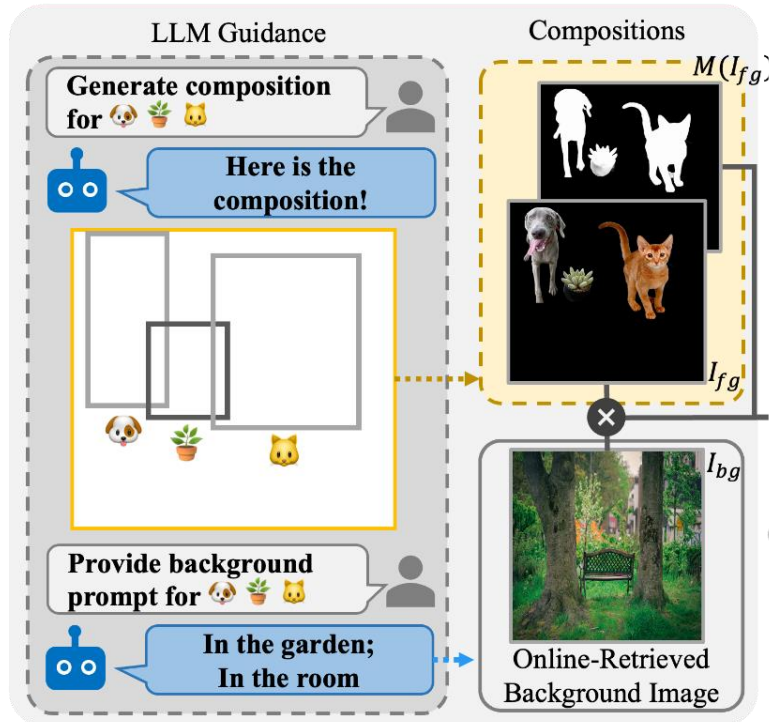
$\mathcal{D}(\text{🐱})$

To Use Generative Models to Create High-Quality Data

1 Object Association & Foreground Segmentation

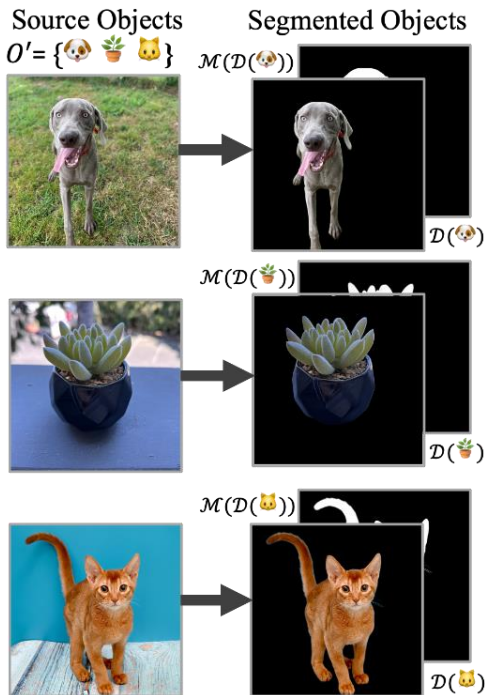


2 LLM-Guided Object Composition

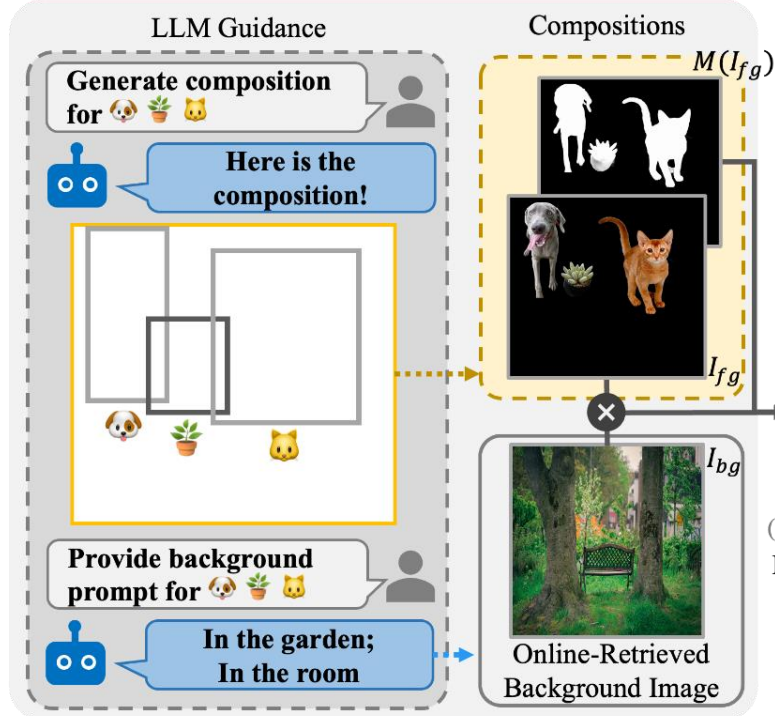


To Use Generative Models to Create High-Quality Data

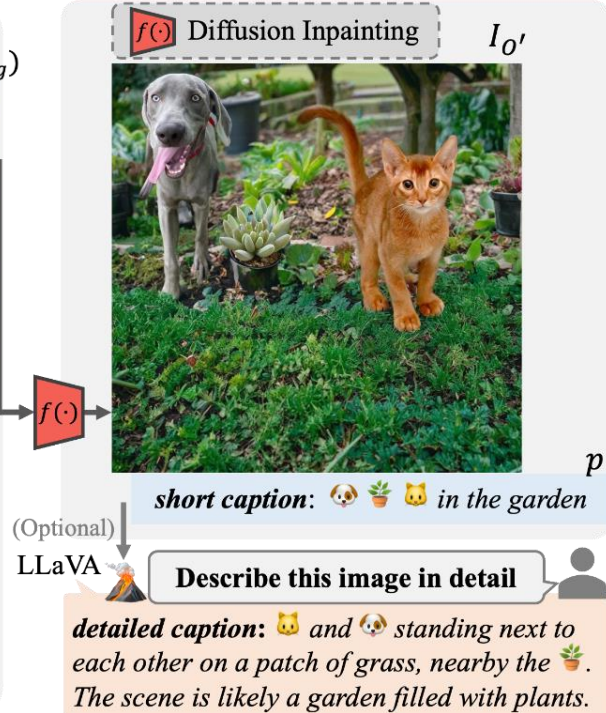
1 Object Association & Foreground Segmentation



2 LLM-Guided Object Composition

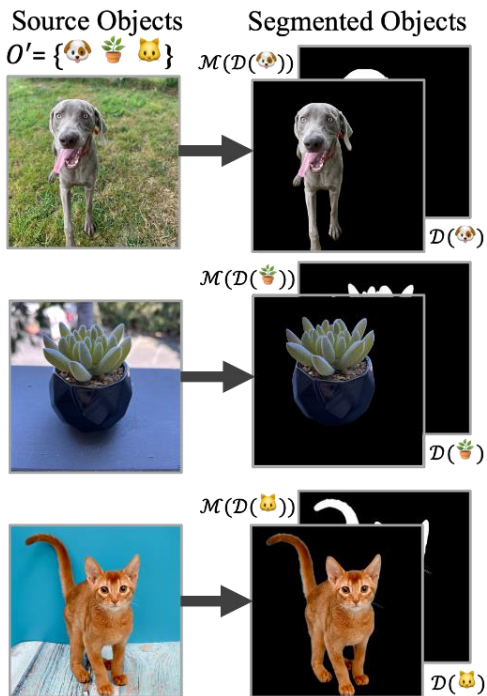


3 Background Repainting & Image Recaptioning



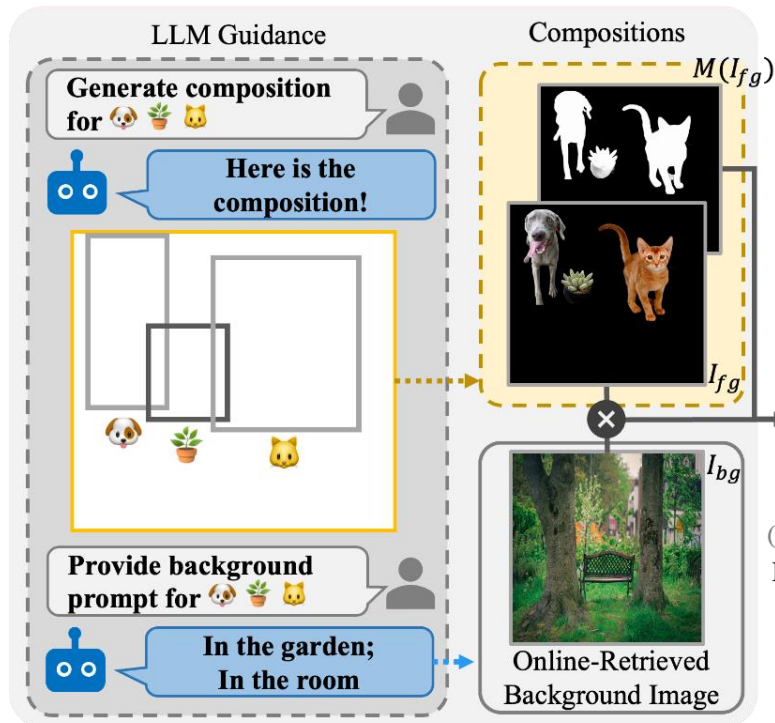
Distillation 2.0?

1 Object Association & Foreground Segmentation



DIS, ECCV'22

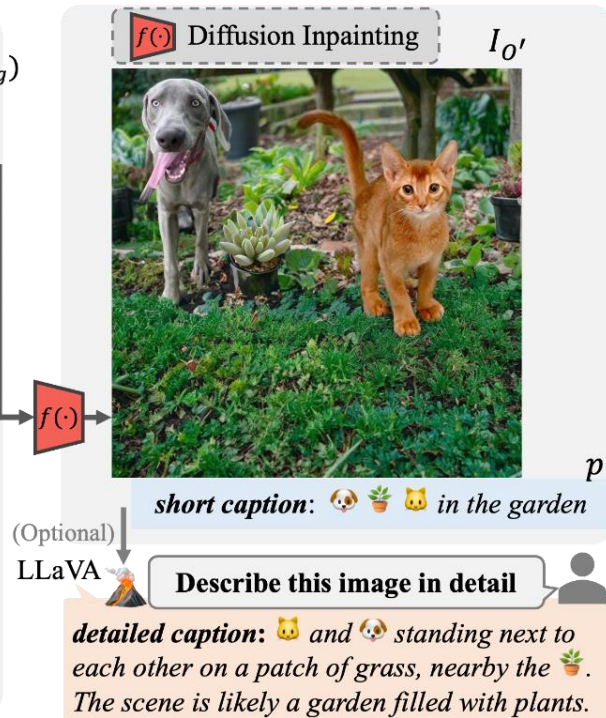
2 LLM-Guided Object Composition



ChatGPT API

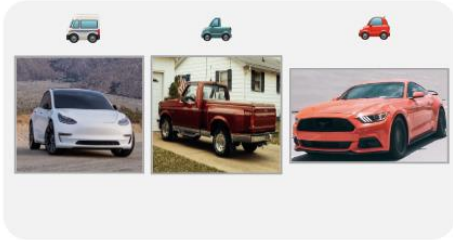
3 Stable Diffusion XL, ICLR'24

3 Background Repainting & Image Recaptioning



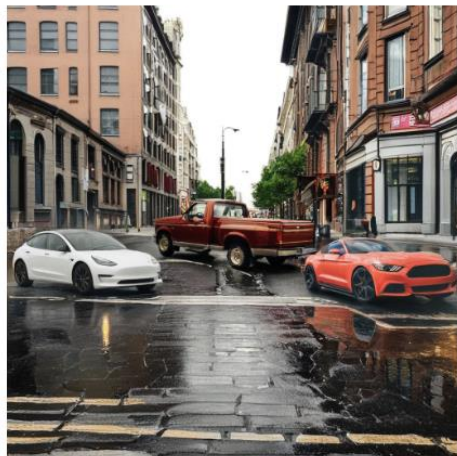
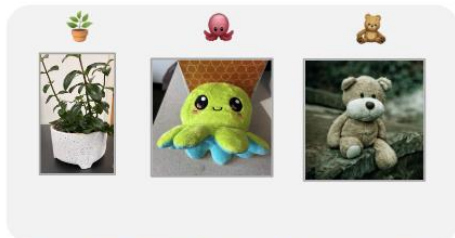
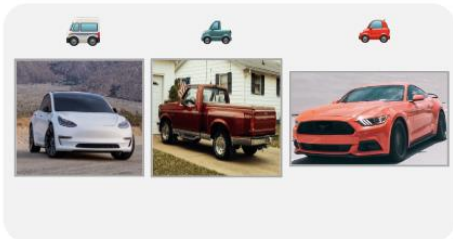
LLaVA, NeurIPS'23

Some Examples from the Generated Dataset

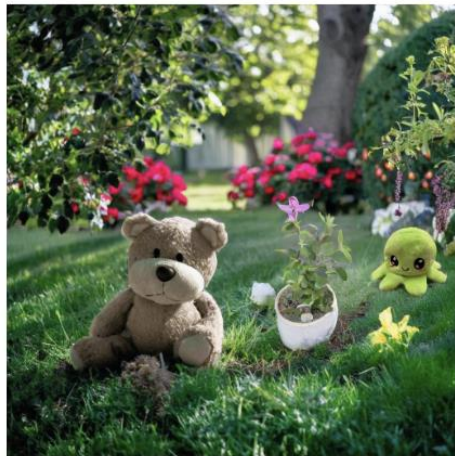


short caption: 🚗 and 🚗 and 🚗 in the city.

Some Examples from the Generated Dataset

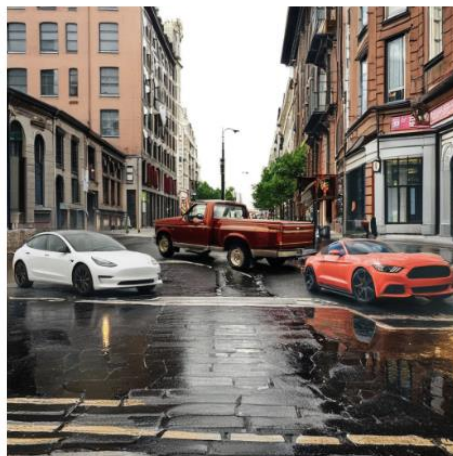
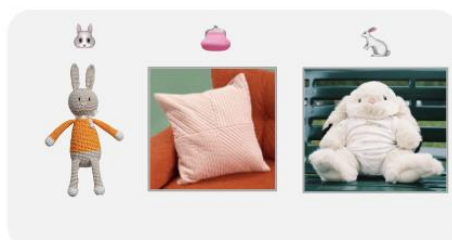
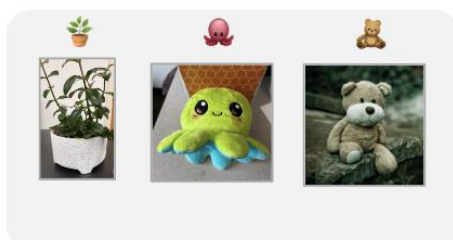
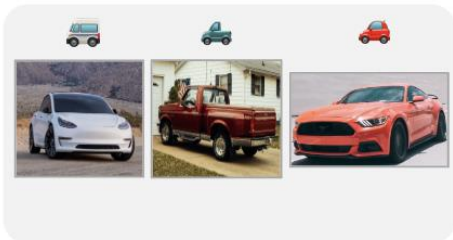


short caption: 🚗 and 🚙 and 🚗 in the city.

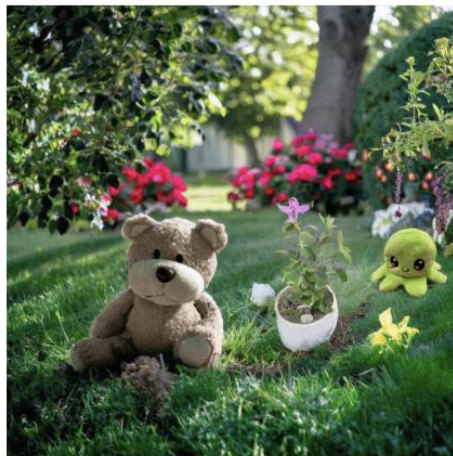


short caption: 🧸 and 🌱 and 🦑 in the garden.

Some Examples from the Generated Dataset



short caption: 🚗 and 🚚 and 🚗 in the city.

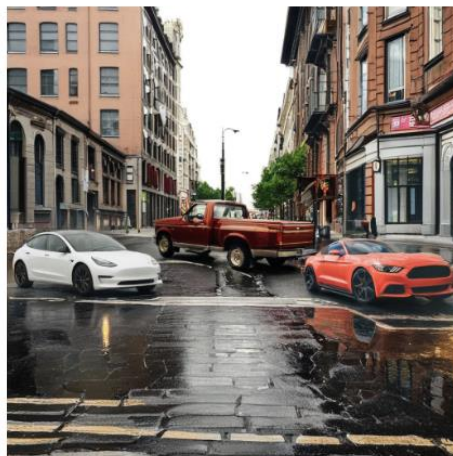
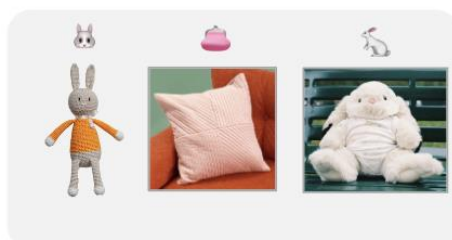
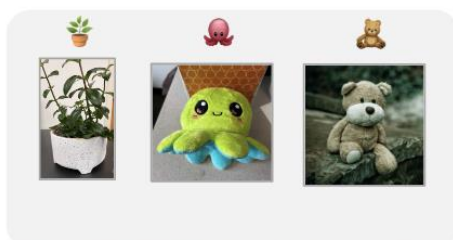
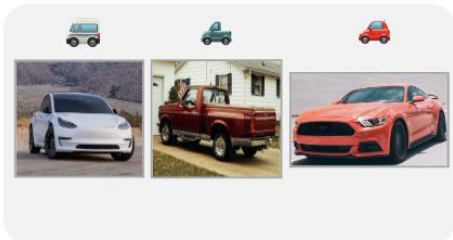


short caption: 🧸 and 🌱 and 🐙 in the garden.

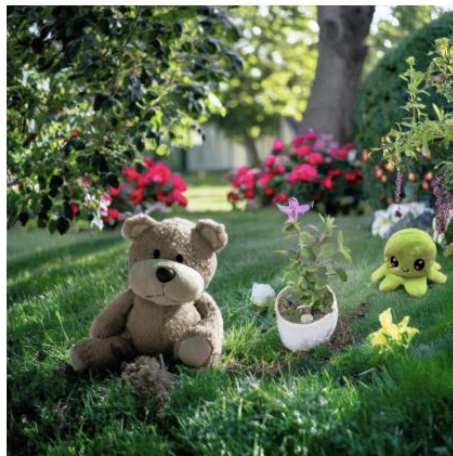


detailed caption: In the playroom, there're 🐰 on the blue cushion, 🐰 on the table, 🛋️ on a white chair, 🚗 on floor, a woven basket, and gray wall.

Some Examples from the Generated Dataset



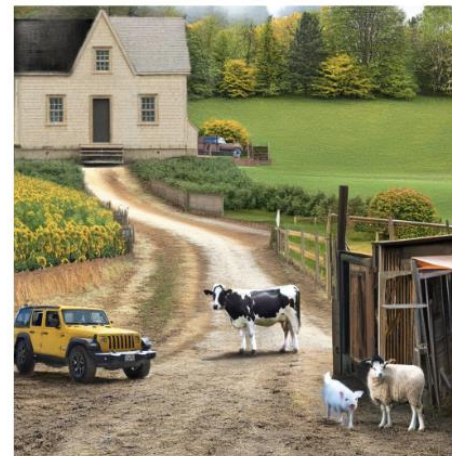
short caption: 🚗 and 🚙 and 🚗 in the city.



short caption: 🧸 and 🌱 and 🍄 in the garden.



detailed caption: In the playroom, there're 🐰 on the blue cushion, 🐰 on the table, 🛋️ on a white chair, toy car on floor, a woven basket, and gray wall.

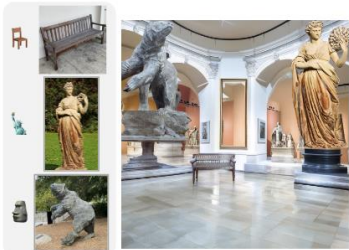


detailed caption: a farm with a 🚙 parked on a dirt road in front of 🏠. 🐷 stands on the road. 🐷 and 🐏 near barn. The 🏠 is surrounded by a green lawn.

Improved Generative Models

Source
Images

MyCanvas



Custom Diffusion + Source Images



Close up view of a 🐱 and a 🐶 hiding in the grass.

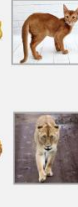

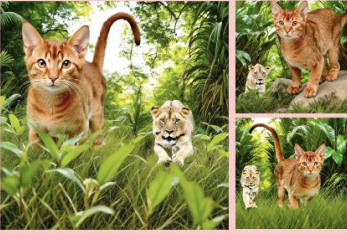














Sunny, autumn, forest, a 🪑 and a 🗑️ on the ground, nearby a 🪑.



The winter season in a sunny day with cloud and lots of snow, the entire ground is all white and christmasy, a 🌲 and a 🏠 and a 🚗 and a 🚙.

Improved Generative Models

Source Images	MyCanvas	Prompting (Ours) + MyCanvas	Custom Diffusion + MyCanvas	Custom Diffusion + Source Images
				
		<p>Close up view of a 🐱 and a 🐾 hiding in the grass.</p>		
				
		<p>Sunny, autumn, forest, a 🦊 and a 🐾 on the ground, nearby a 🪑.</p>		
				
		<p>The winter season in a sunny day with cloud and lots of snow, the entire ground is all white and christmasy, a 🏠 and a 🍷 and a 🚗 and a 🚙.</p>		

Using Generative Model to Improve Any Model

- Sim2Real
- Defect Synthesis for Anomaly Detection
- Virtual Speakers
- Generative Acoustic Simulation
- ...
- Data-Centric efficient AI

Main Points

- 1. **Generative models can be cascaded** to form a data pipeline
- 2. **Generated data can improve generative models** in return
- 3. We can use this data-centric method to **improve any model**

Yubei Chen, ybchen@ucdavis.edu
Assistant Professor
ECE Department, UC Davis
Cofounder, Aizip.ai

