Stylebreeder: Exploring and Democratizing Artistic Styles through Text-to-Image Models

Matthew Zheng*, Enis Simsar*, Hidir Yesiltepe, Federico Tombari, Joel Simon, Pinar Yanardag

VIRGINIA TECH · Google · ETH zürich · Artbreeder · NEURAL INFORMATION PROCESSING SYSTEMS

Introduction

Recent AI-generated datasets contain limited stylistic diversity and skew toward user groups, reflecting constrained environments and time frames

- We introduce a novel extensive AI-generated dataset Stylebreeder
- We identify unique image styles in an unsupervised manner
- We showcase a recommendation system that aligns style suggestions with individual preferences
- We provide public access to download pre-trained style LoRAs for personalized content generation through our web platform Style Atlas

Stylebreeder Dataset

| Image	Prompts	Additional Features
	Positive Prompt black and red and gold shot of a meadow in spring, flowers everywhere, lush green grass, beautiful sky, realistic faces, girl in a flowy dress and flowers in her arms, dark fantasy, intricate details, hyper detailed, jean baptiste monge, carne griffiths, michael garmash, tsutomu nihei, motifs of spring and renewal : 1 **Negative Prompt** bad hands, amateur, low quality, mangled hands, disfigured hands, ugly hands : -0.5	**ImageID** 298709787, **UserID** 174769, **Identity Attack** 0.00456, **Prompt NSFW** 0.00015; **Model** sdxl-1.0, **Step** 13, **Insult** 0.0031, **Image NSFW** 0.00573; **Image Size** (640, 960), **Seed** 556, **Threat** 0.001, **Toxicity** 0.00364; **Timestamp** 2024-03-20 3:02:23, **CFG Scale** 26, **Cluster ID** 3398, **Obscene** 0.00037

- 6.8 million images generated by 95,000 unique users scraped from the Artbreeder website from July 2022 to May 2024
- It includes detailed metadata such as Positive Prompt, Negative Prompt, UserID, Timestamp, and Image Size
- We supply model-related hyperparameters, including Model Type, Seed, Step, and CFG Scale
- Further metadata like Cluster ID, along with scores for Prompt NSFW, Image NSFW, and Toxicity computed using state-of-the-art models

Name	Source	# Images	Year	Original Prompt Included	Multiple Models
DiffusionDB	SD Discord	14,000,000	Aug 2022	✓	✗
Midjourney Kaggle	Midjourney	250,000	Jun 2022-Jul 2022	✓	✗
TWIGMA	Twitter	800,000	Jan 2020-Mar 2023	✗	✓
STYLEBREEDER (Ours)	Artbreeder	6,818,217	Jul 2022-May 2024	✓	✓

(a) NSFW scores for images   (b) NSFW and other scores for text prompts
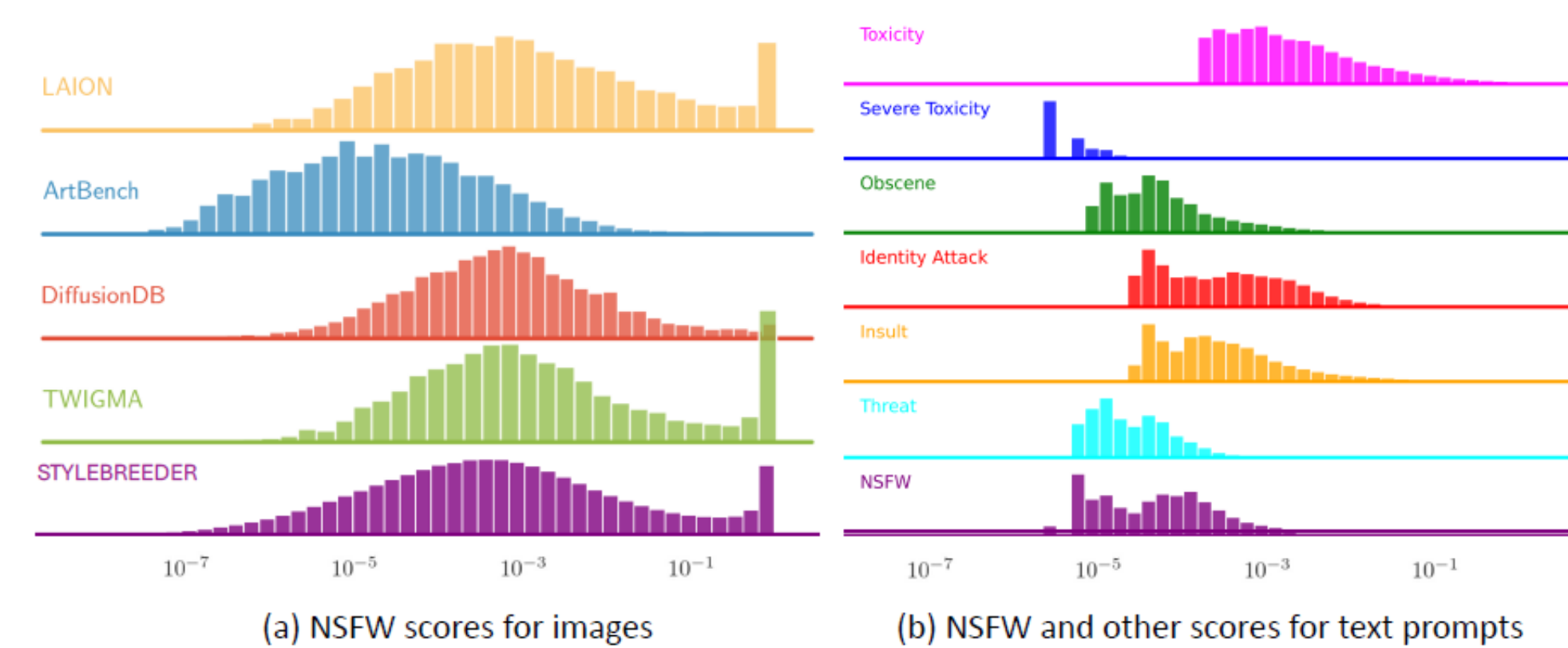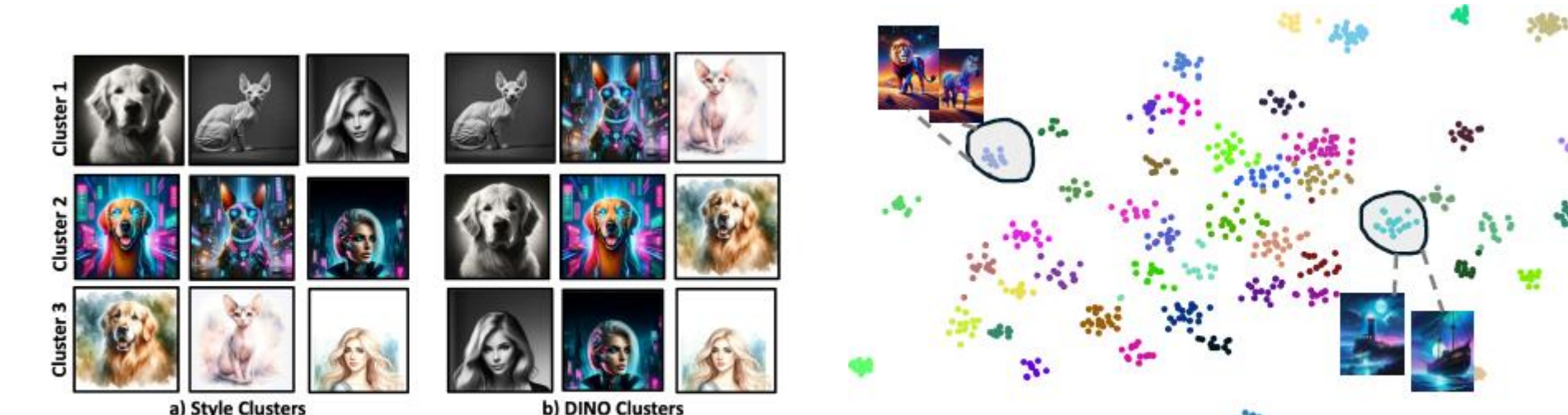
Discovering Diverse Artistic Styles

- We convert the images into a set of style embeddings using a state-of-the-art feature extractor, CSD
- These embeddings are then clustered into groups using the K-Means++ algorithm utilizing cosine similarity to ensure cohesion within clusters
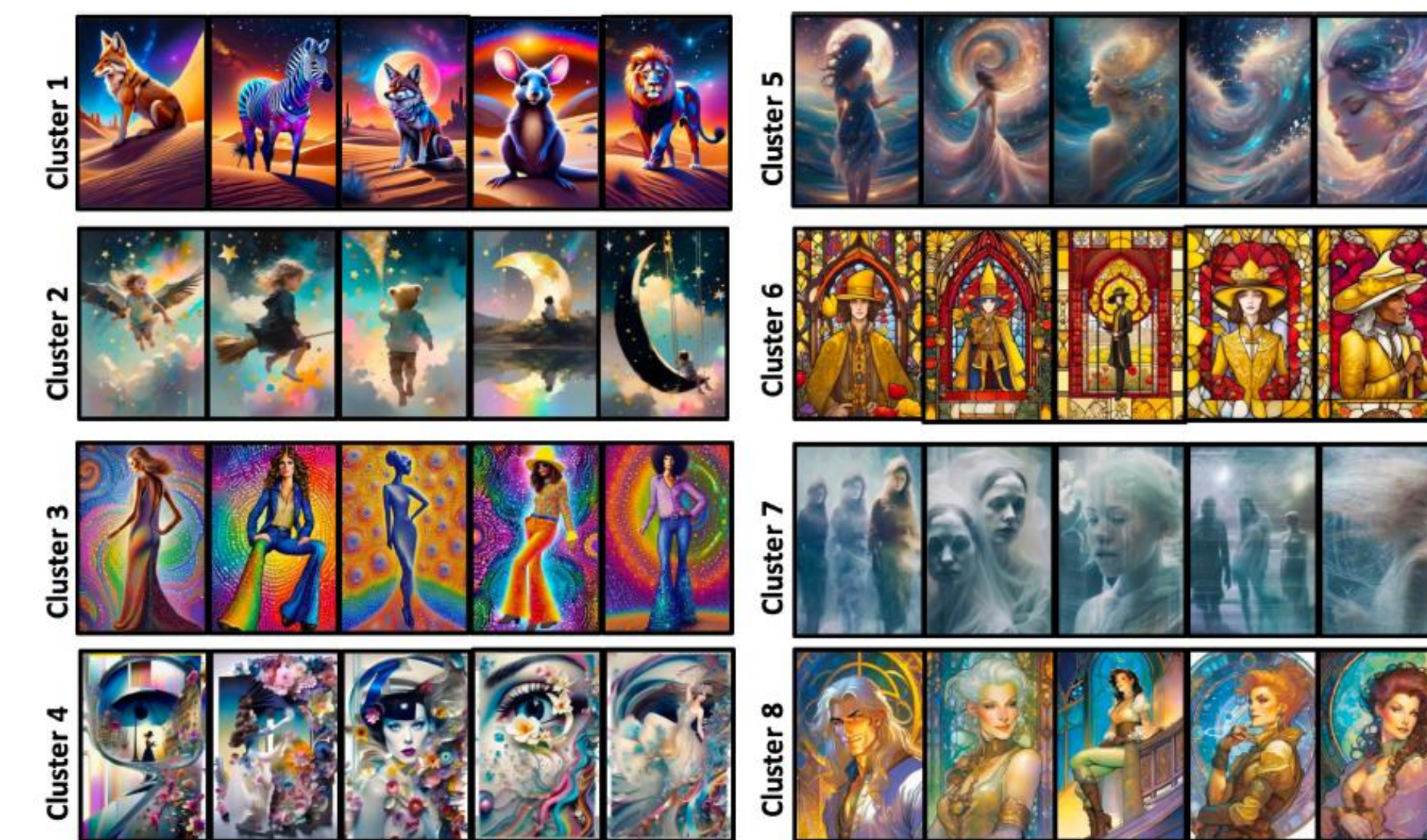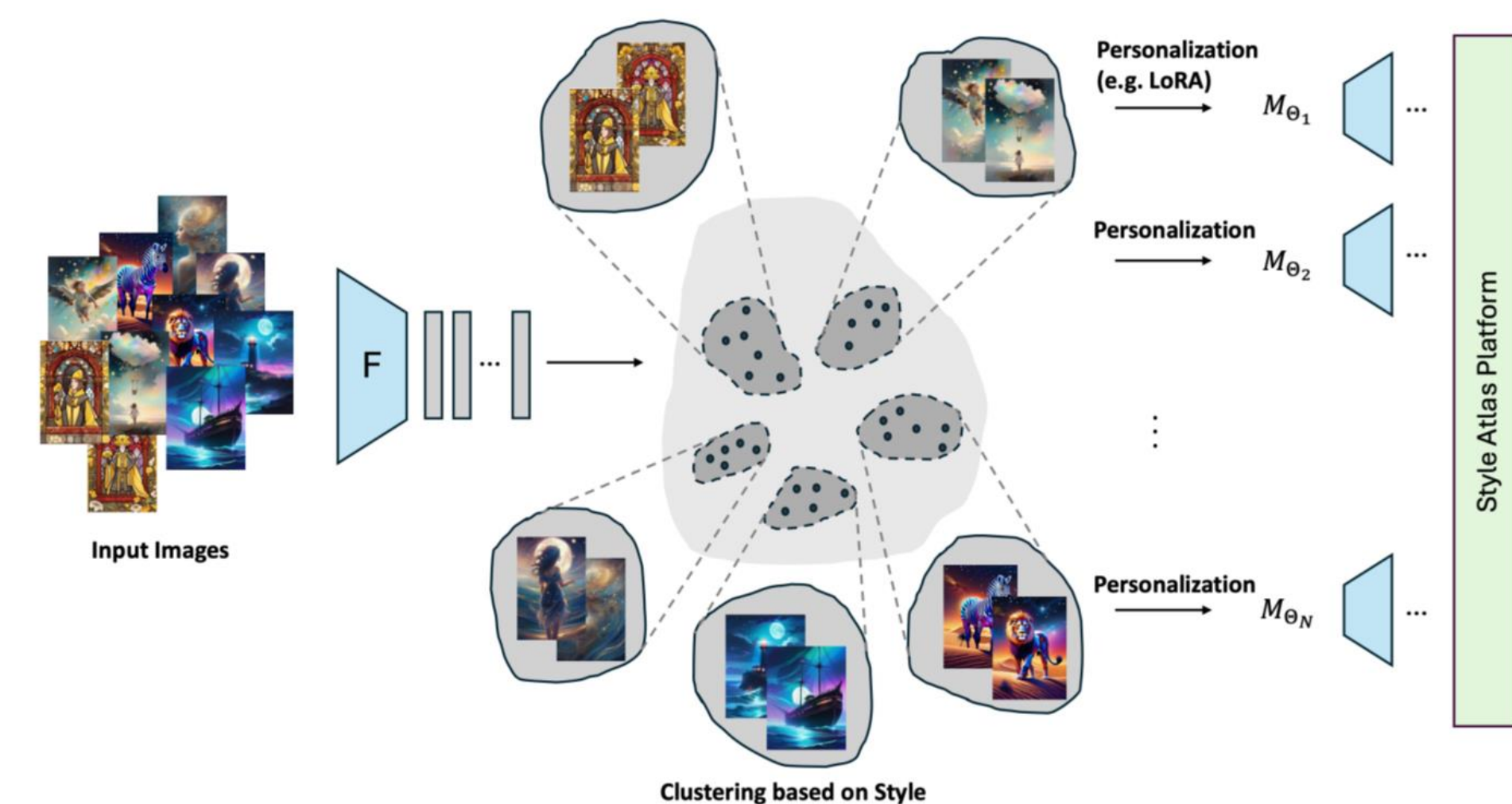

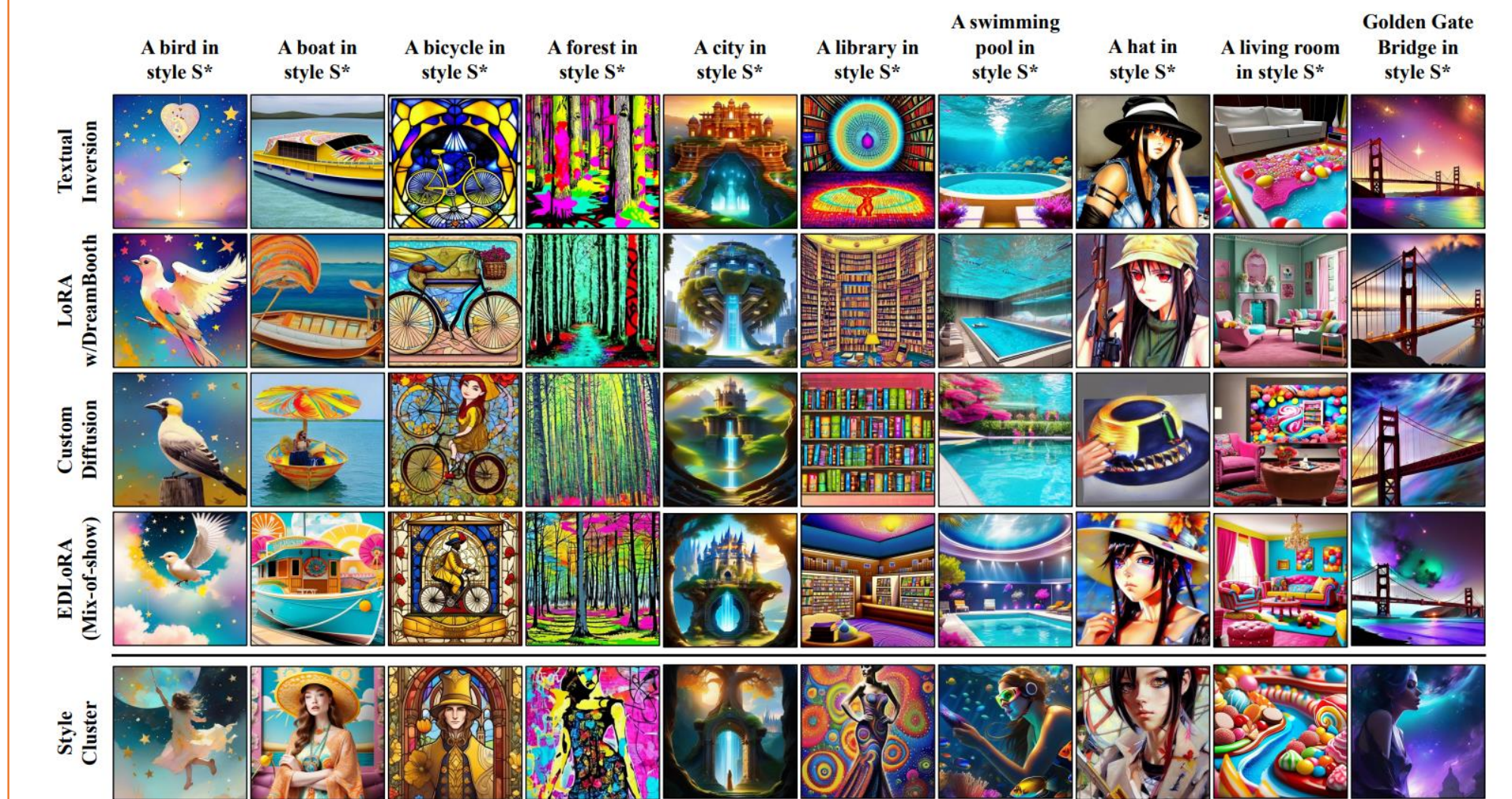(a) User-generated images from 10 random clusters

(b) Style-based vs. traditional clustering   (c) t-SNE plot of 50 random clusters

We use these clusters of images of discovered styles to fine-tune LoRA models capable of generating new images with similar styles

Clustering based on Style

Experiments

		Textual Inversion	LoRA w/DreamBooth	EDLoRA	Custom-Diffusion
CLIP-I	Avg.	0.6869 ± 0.10	0.6299 ± 0.11	$\mathbf{0.6957 \pm 0.11}$	0.5917 ± 0.12
	Min.	0.6166 ± 0.11	0.5654 ± 0.11	$\mathbf{0.6214 \pm 0.11}$	0.5324 ± 0.11
	Max.	0.7428 ± 0.10	0.6831 ± 0.11	$\mathbf{0.7521 \pm 0.12}$	0.6440 ± 0.12
CLIP-T	Avg.	0.1857 ± 0.02	$\mathbf{0.1896 \pm 0.02}$	0.1822 ± 0.01	0.1809 ± 0.02
	Min.	0.1555 ± 0.02	$\mathbf{0.1573 \pm 0.02}$	0.1527 ± 0.01	0.1486 ± 0.02
	Max.	0.2392 ± 0.03	$\mathbf{0.2663 \pm 0.03}$	0.2389 ± 0.03	0.2585 ± 0.03
DINO	Avg.	0.3801 ± 0.15	0.2668 ± 0.17	$\mathbf{0.4125 \pm 0.18}$	0.2546 ± 0.17
	Min.	0.2581 ± 0.13	0.1682 ± 0.14	$\mathbf{0.2790 \pm 0.15}$	0.1634 ± 0.14
	Max.	0.4838 ± 0.17	0.3585 ± 0.19	$\mathbf{0.5246 \pm 0.19}$	0.3402 ± 0.19

Style-based Recommendation

- We use a matrix-factorization approach, which involves a set of items where users rate items they have interacted with
- Users are the creators who generate images, and items are the clusters in which generated images are assigned

Style Atlas

We provide 100 style LoRAs on our Style Atlas platform so users can browse and download LoRA models for appealing styles