그림 그리는 AI

Zero-Shot Text-to-Image Generation

DALL-E

GDSC CAU OMS

오송경

CONTENTS

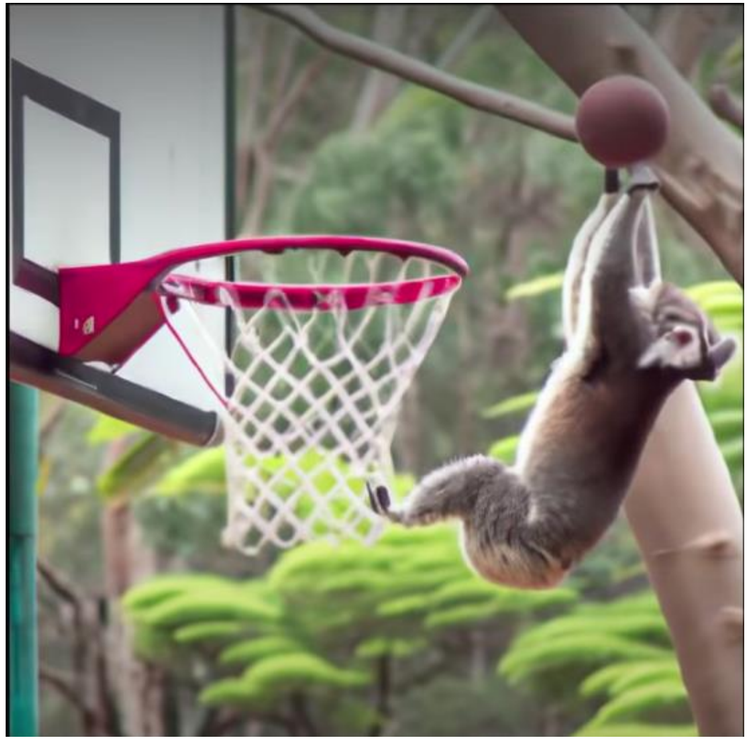
01 DALL-E 란?

02 DALL - E 사용 예시

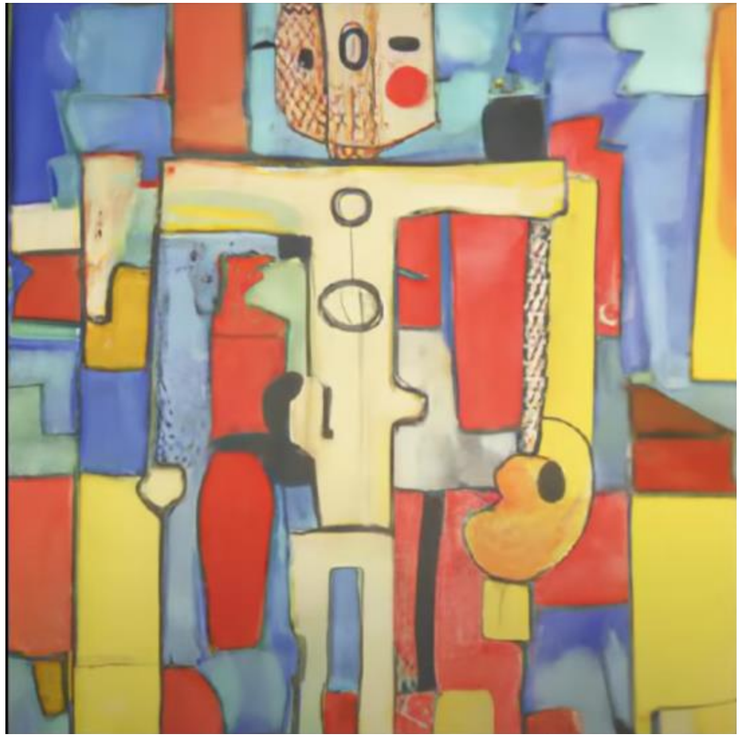
03 DALL - E 기술

01 DALL-E 사용예시

기타치는 북극곰



농구하는 여우원숭이



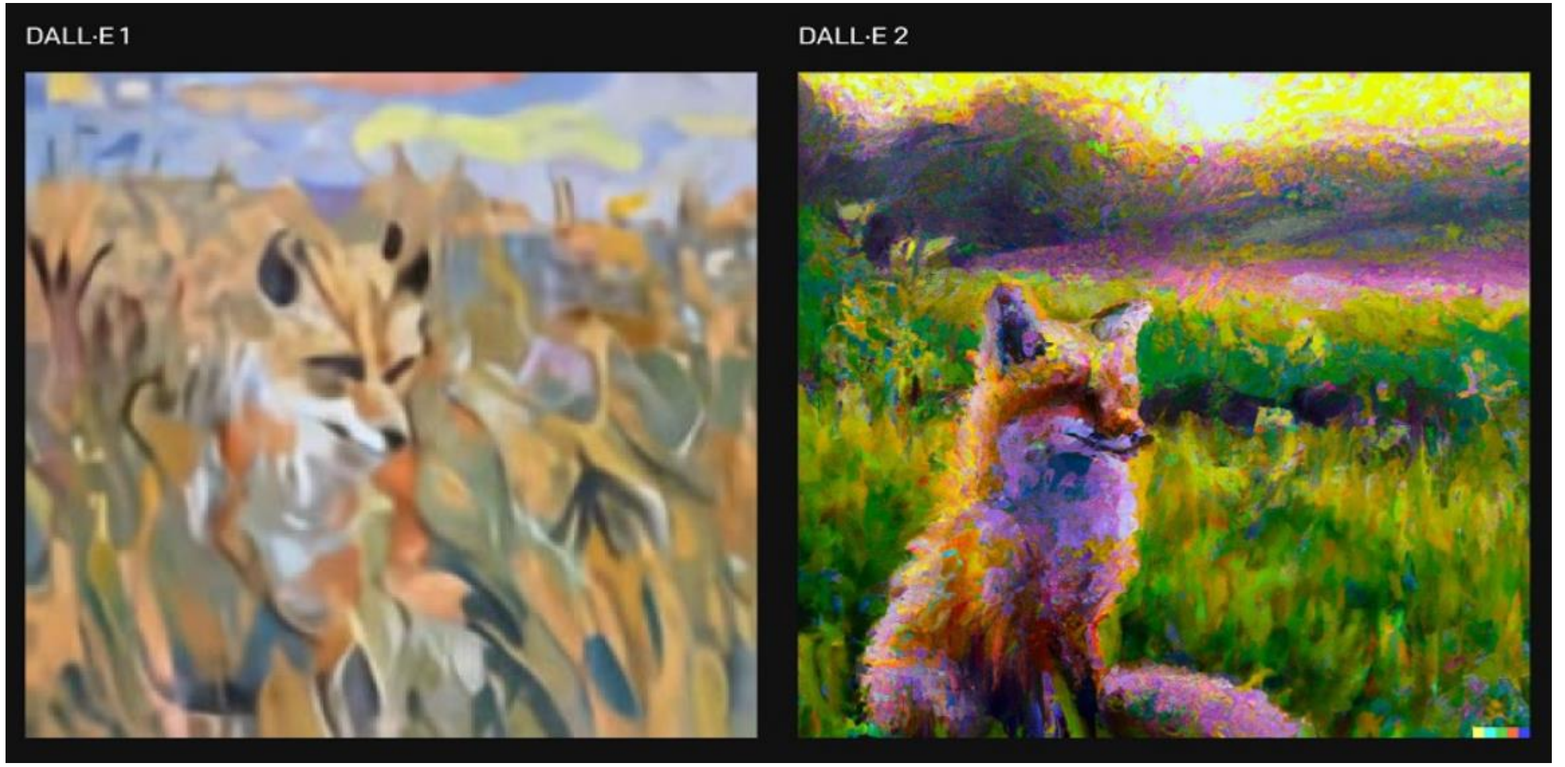
피카소 화풍의 로봇

01 DALL-E란?

- 오픈 ai의 이미지 생성 인공지능 기술
- 자연어처리 + 컴퓨터 비전 기술을 이용해 텍스트에서 이미지를 생성
- 이미지 대상 명칭과 이미지를 연결하는 것이 아닌, 학습 데이터를 조합해 새로 만들어 냄
- GPT-3와 같은 auto regressive transformer 기반

가장 큰 특징은 실제 존재하지 않는
이미지를 만들어낸다는 것!

02 DALL-E 2



2021.01 DALL-E

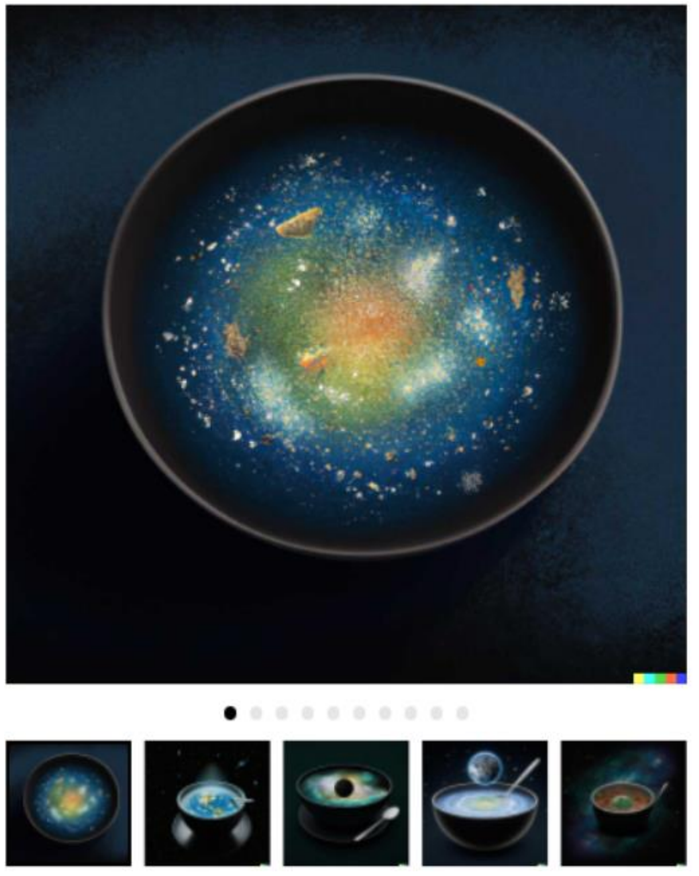
2022.01 DALL-E 2
더 높은 해상도,
더 빠른 속도

02 DALL-E 2

https://openai.com/dall-e-2/

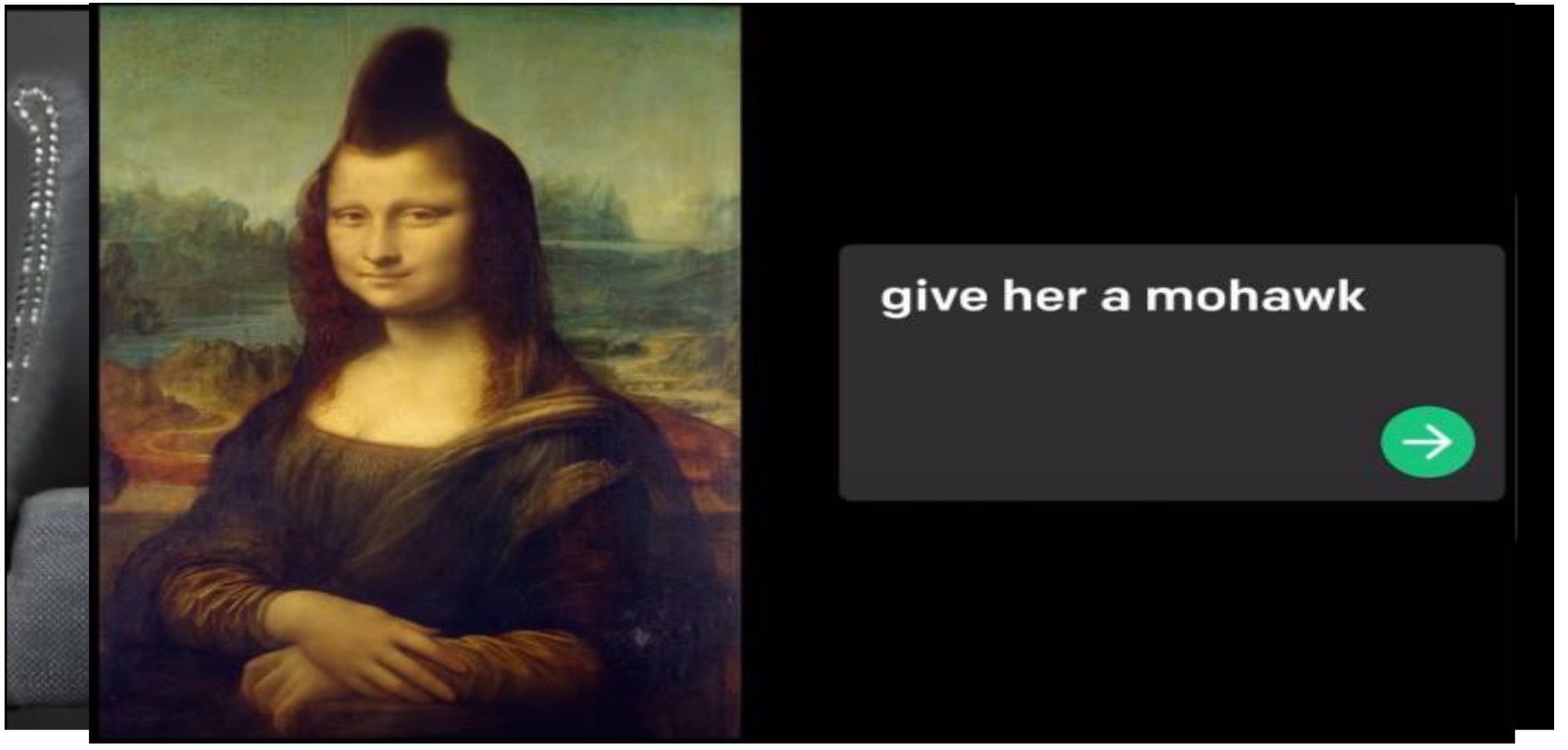
An astronaut riding a horse

Teddy bears shopping for groceries



A bowl of soup as a planet in the universe

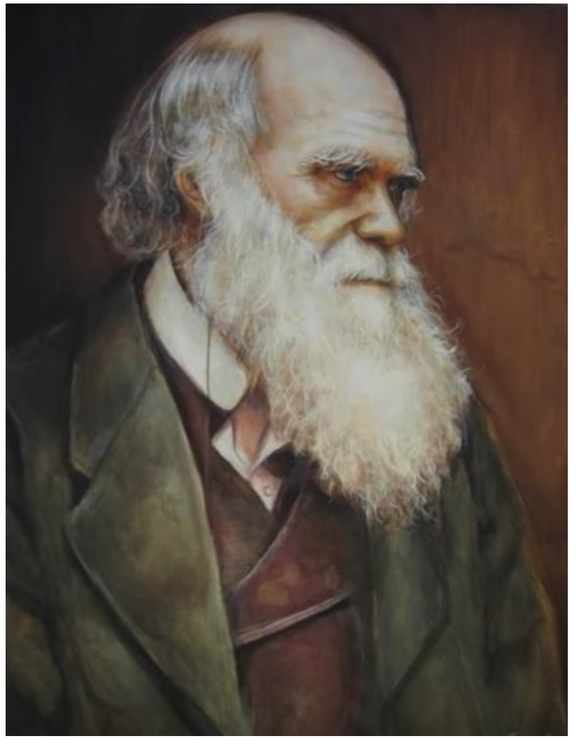
02 DALL-E 2



in-painting : 현실적인 수정과 리터치 기법

맥락을 반영한 채우기

02 DALL-E 2



변형(variation): 각도, 화풍 변화

02 DALL-E 2

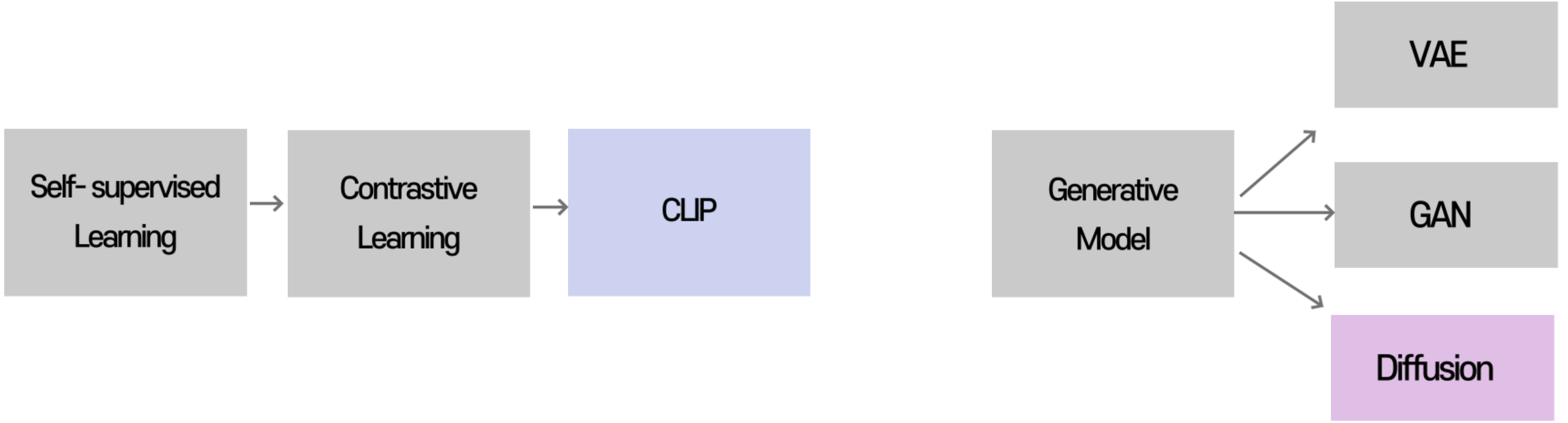
a photo of a cat → an anime drawing of a super saiyan cat, artstation

a photo of a victorian house → a photo of a modern house

language-guided image manipulation

03 DALL-E의 기술

CLIP + DIFFUSION

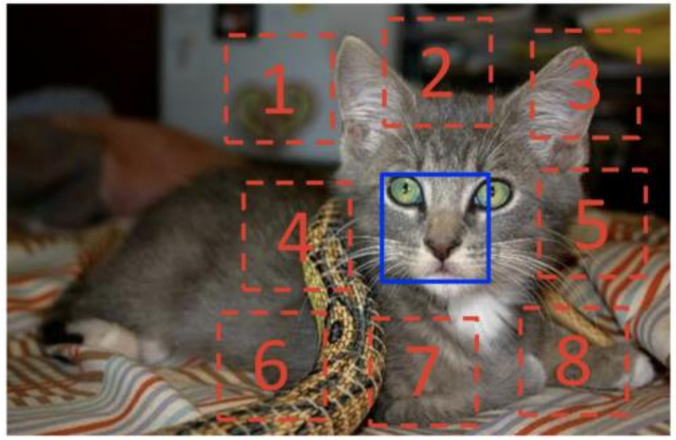


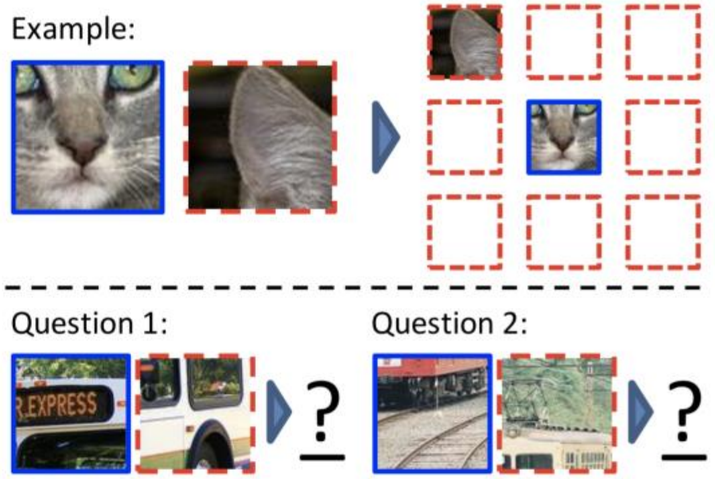
CLIP (Contrastive Language-Image Pre-training)

03 DALL-E의 기술- CLIP

Self-Supervised Learning (자기지도학습)

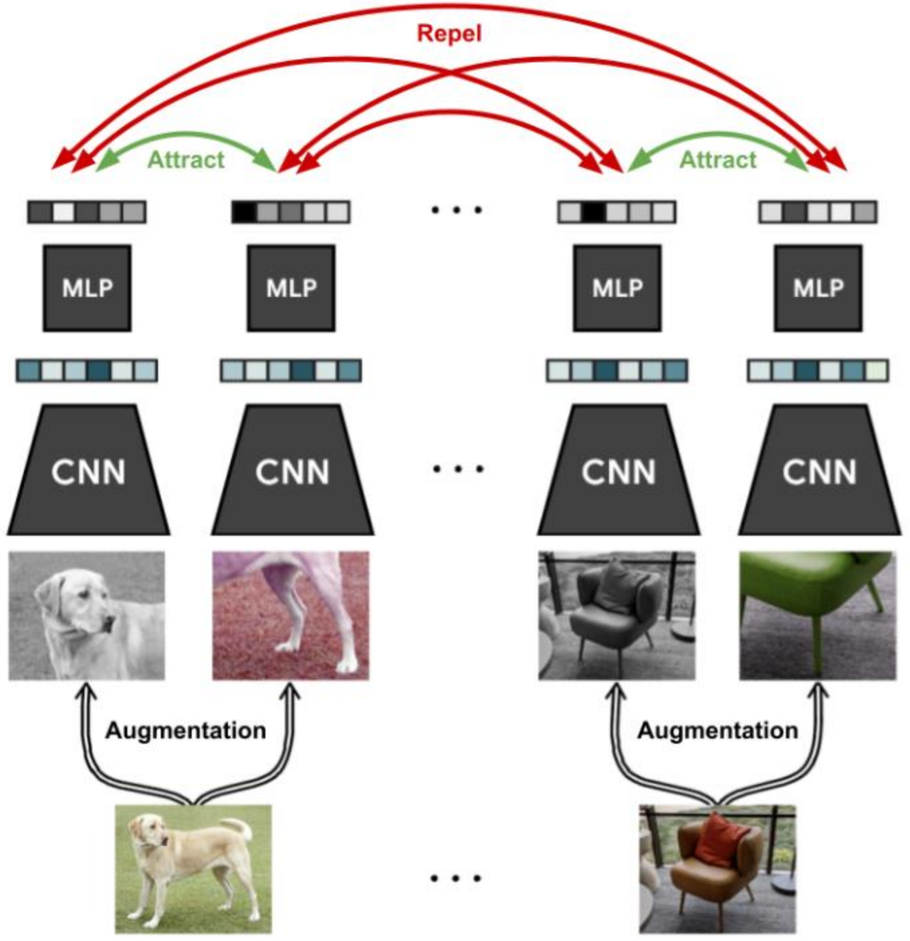
- Unsupervised learning의 한 종류
- 라벨을 만들어서 pre-training



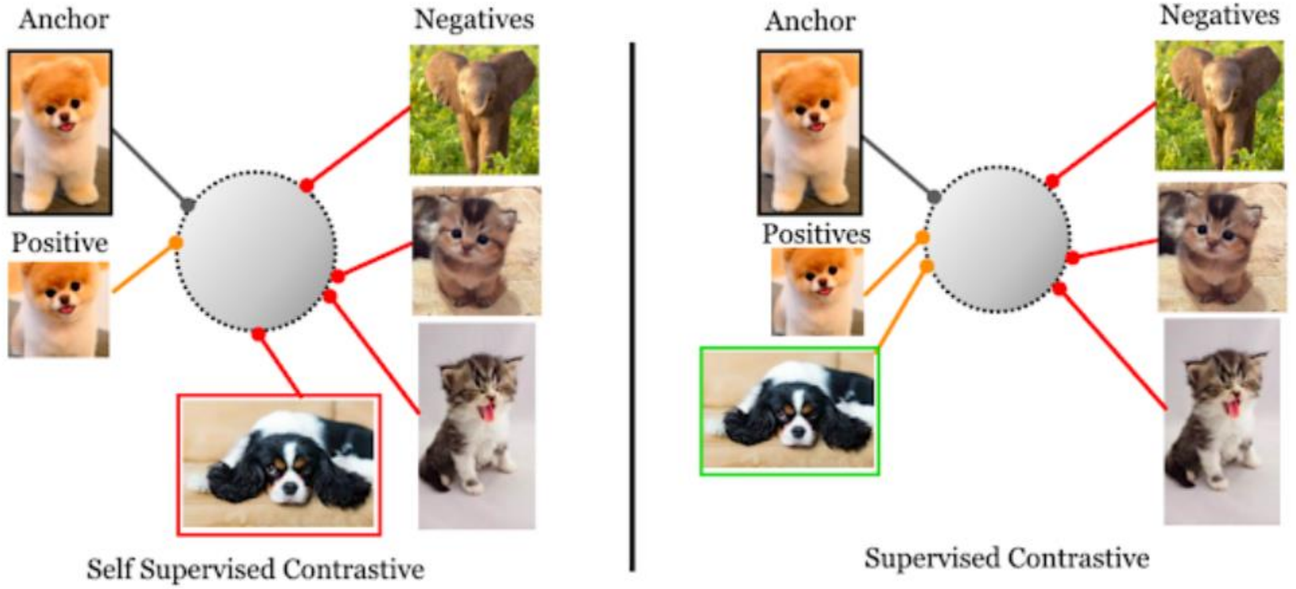


03 DALL-E의 기술- CLIP

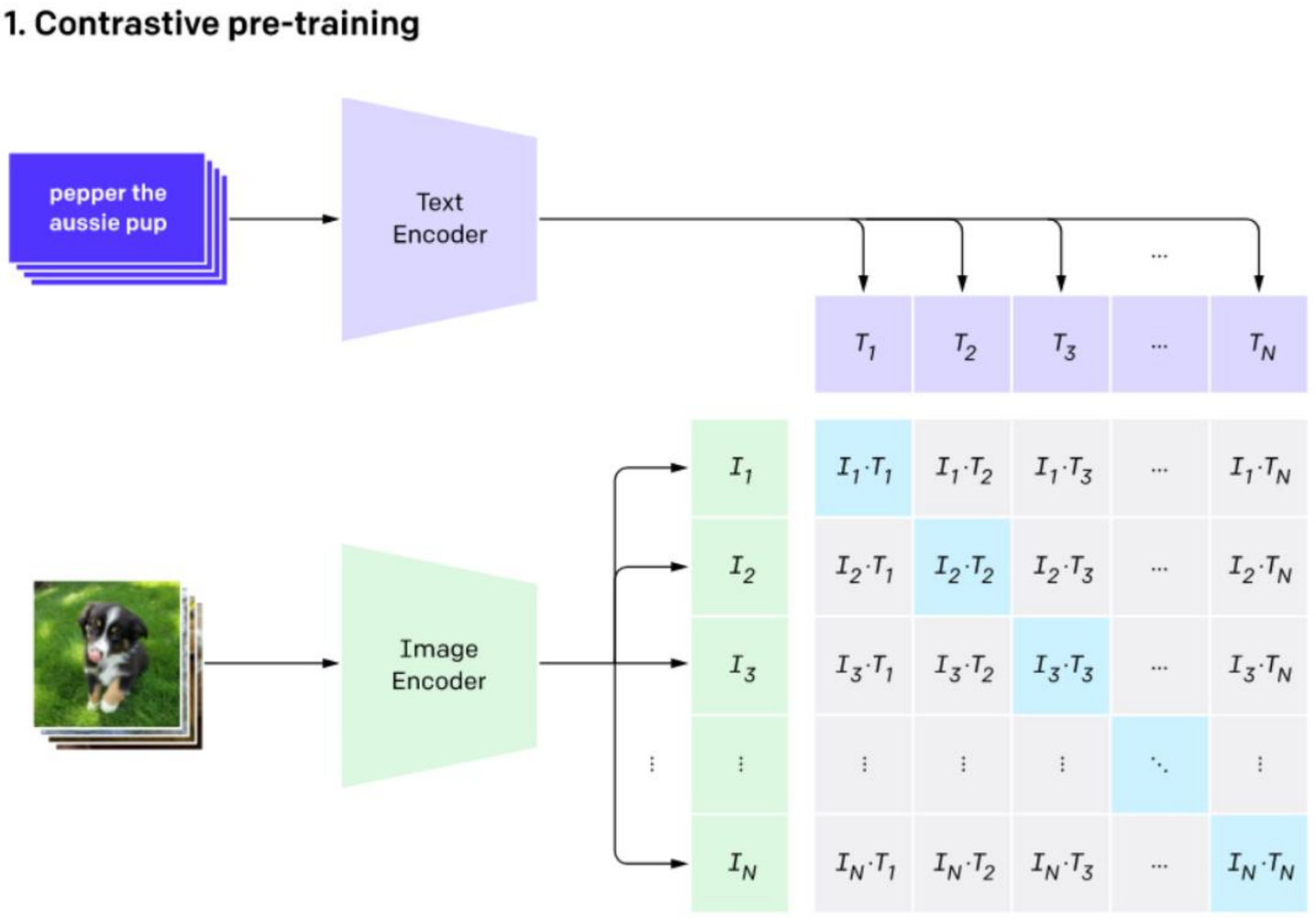
Contrastive Learning



Supervised Contrastive Learning



03 DALL-E의 기술- CLIP



03 DALL-E의 기술- CLIP

Conceptual Caption data set

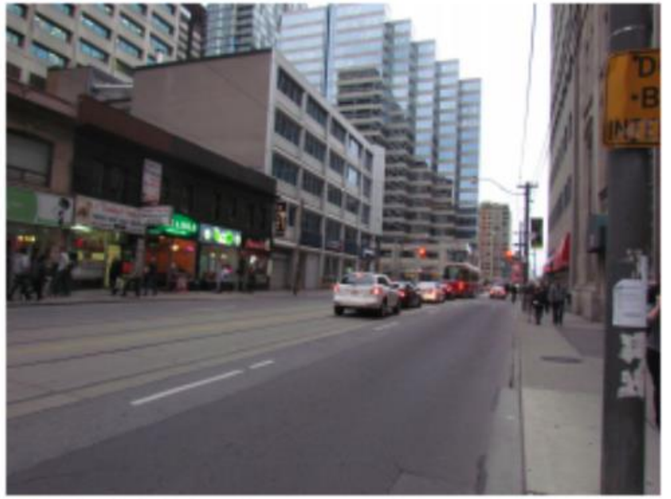
"trees in a winter snowstorm"

"a cartoon illustration of a bear waving and smiling"

"the scenic route through mountain range includes these unbelievably coloured mountains"

"facade of an old shop"

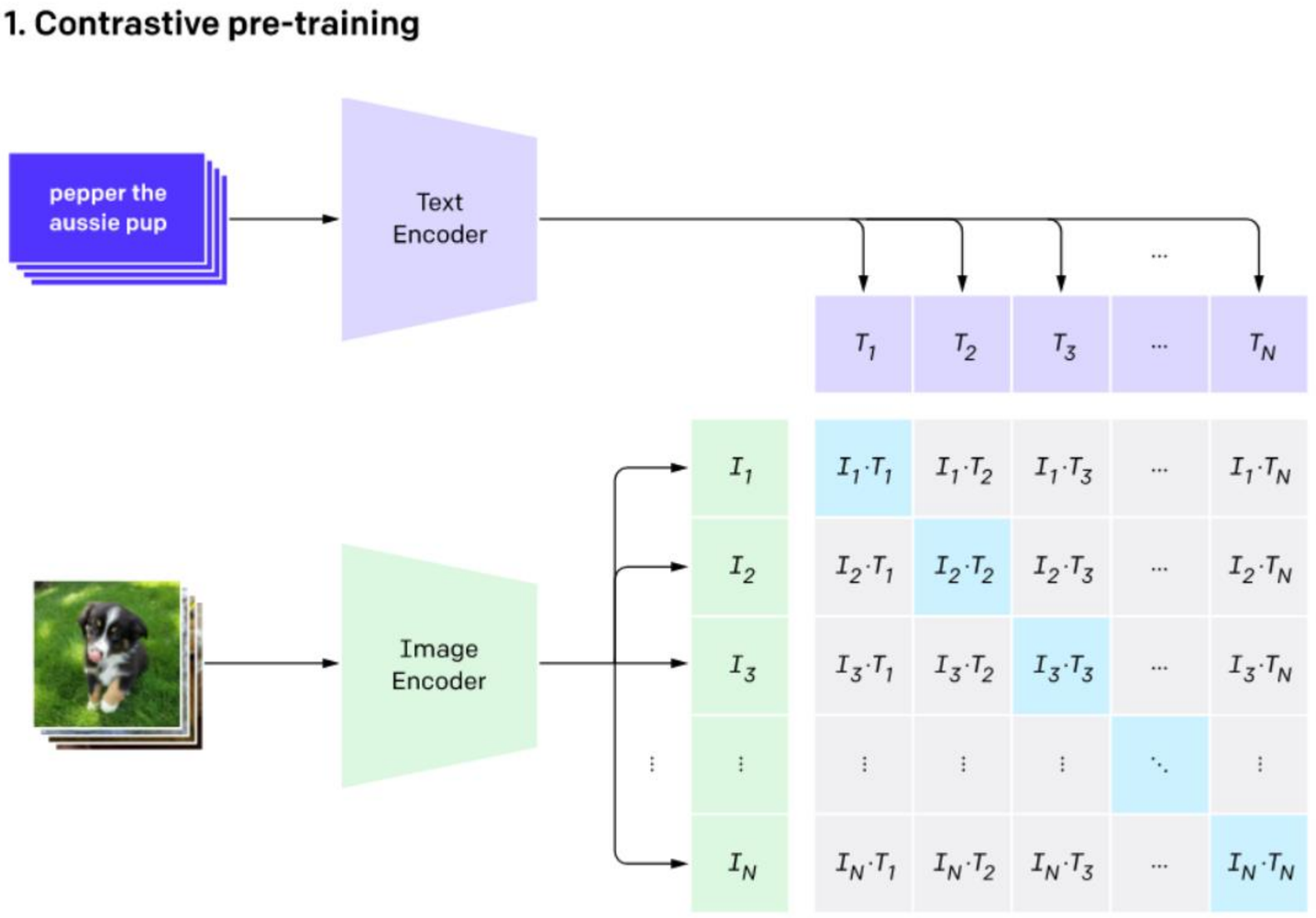
YFCC100M



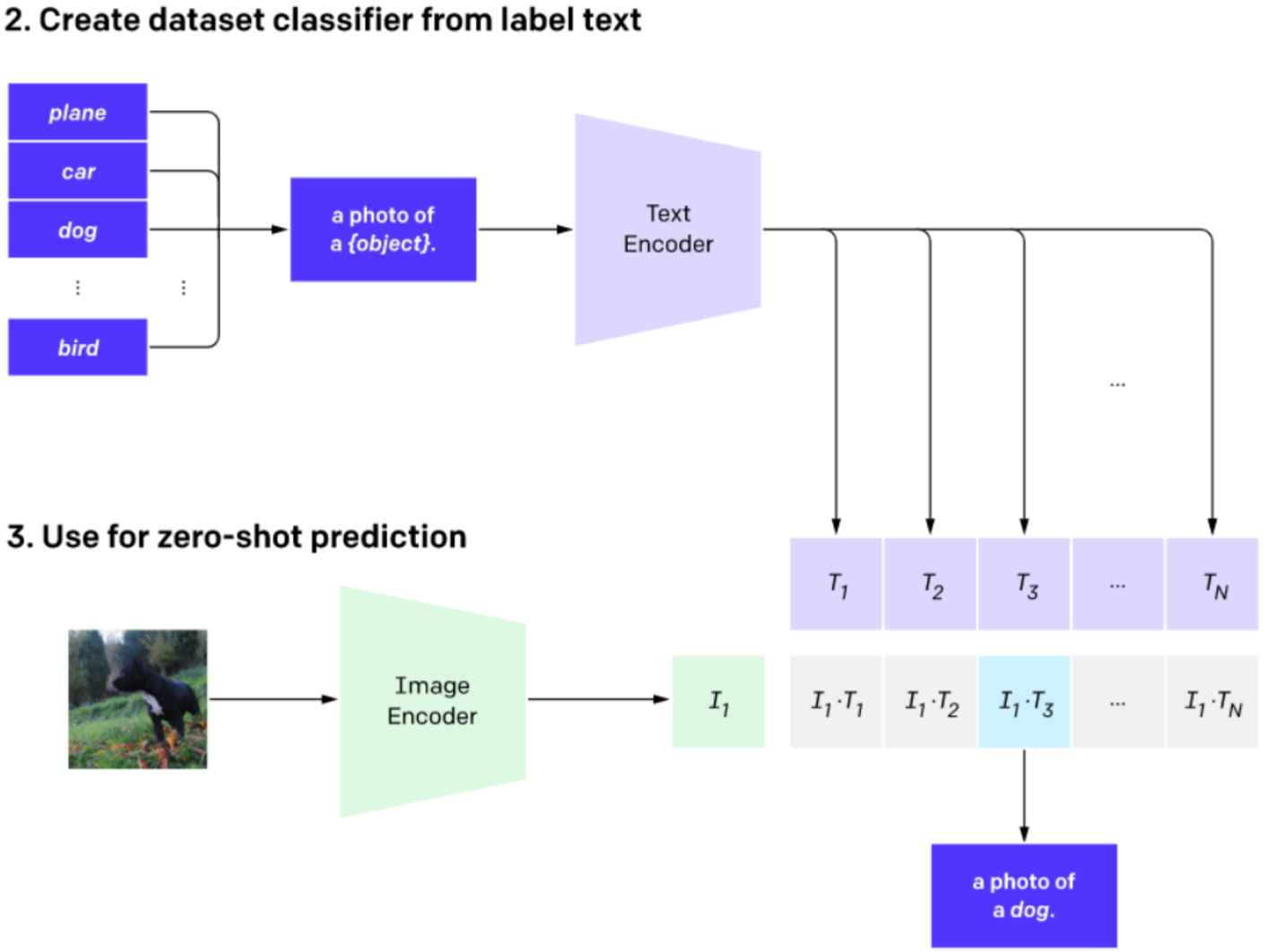
(a) *IMG_9793: Streetcar (Toronto Transit)* by Andy Nystrom CC BY NC ND https://flic.kr/p/jciMdz.

(b) *Celebrating our 6th wedding anniversary in Villa Mary* by Rita & Tomek CC BY NC SA https://flic.kr/p/fCXEJi.

03 DALL-E의 기술- CLIP



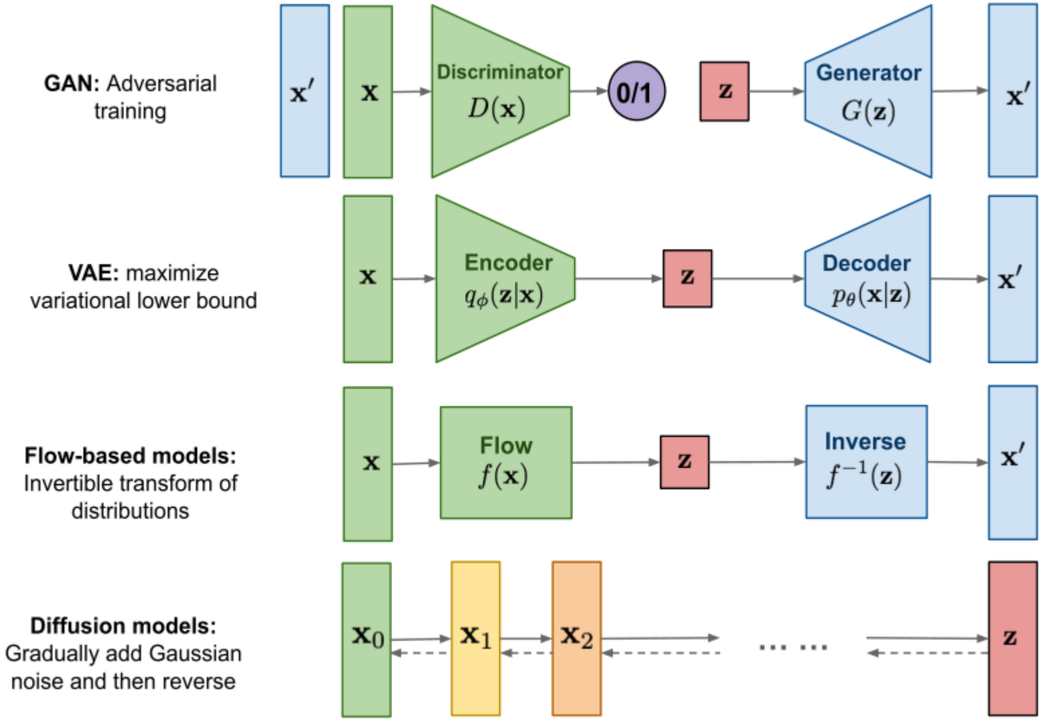
03 DALL-E의 기술- CLIP

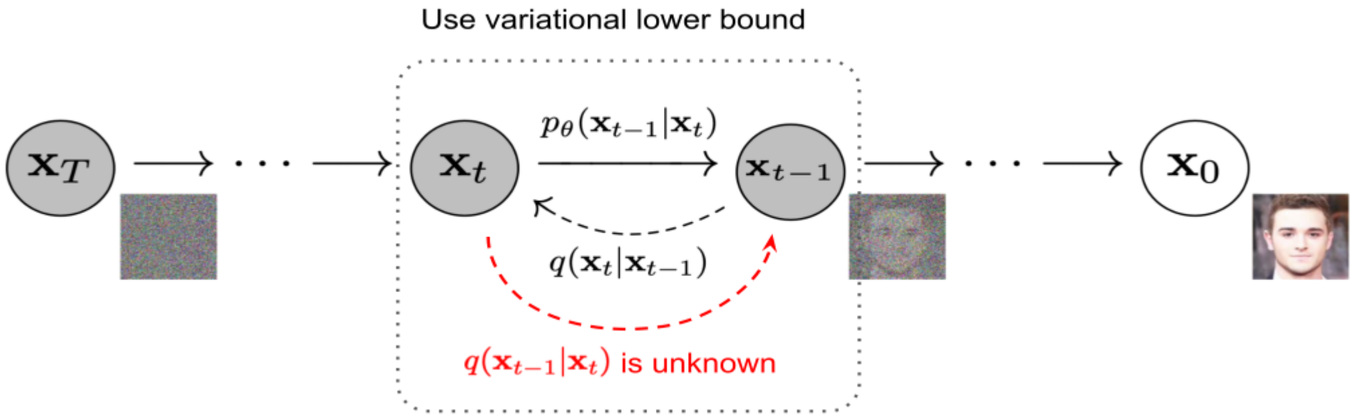


03 DALL-E의 기술- CLIP

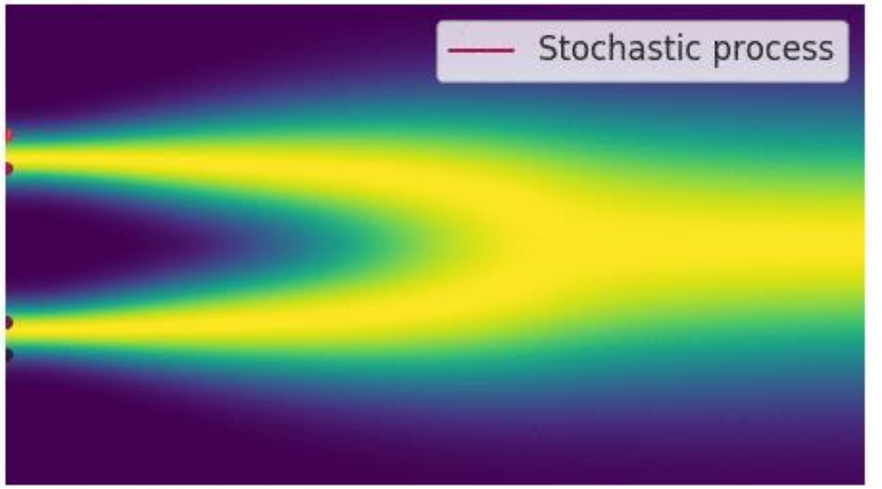
DATASET	IMAGENET RESNET101	CLIP VIT-L
 ImageNet	76.2%	76.2%
 ImageNet V2	64.3%	70.1%
 ImageNet Rendition	37.7%	88.9%
 ObjectNet	32.6%	72.3%
 ImageNet Sketch	25.2%	60.2%
 ImageNet Adversarial	2.7%	77.1%

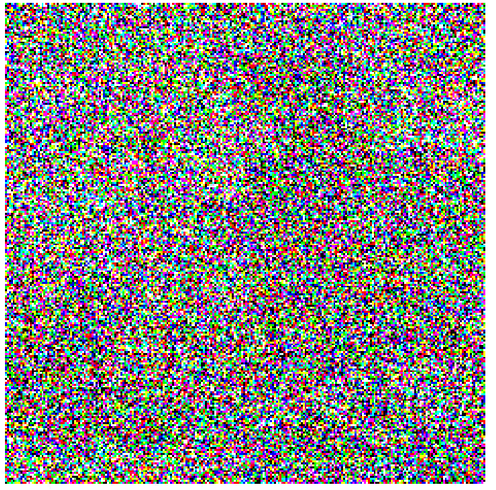
03 DALL-E의 기술- Diffusion model

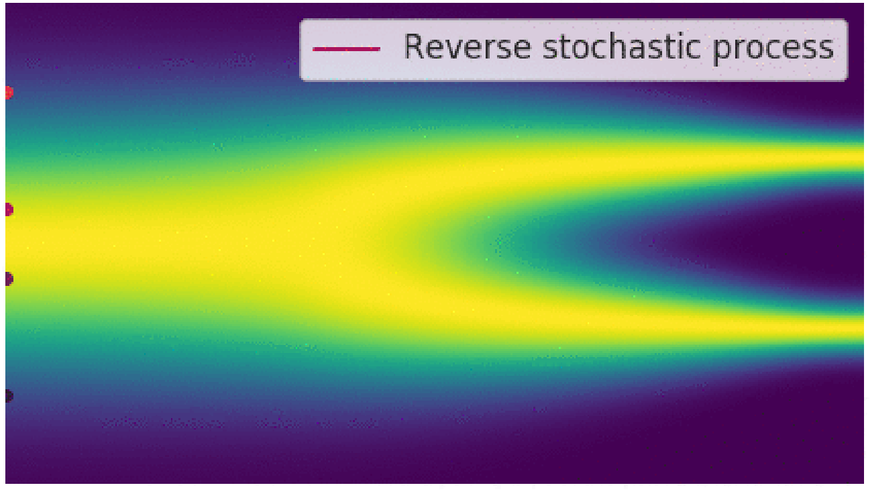




03 DALL-E의 기술- Diffusion model







03 DALL-E의 기술- Diffusion model

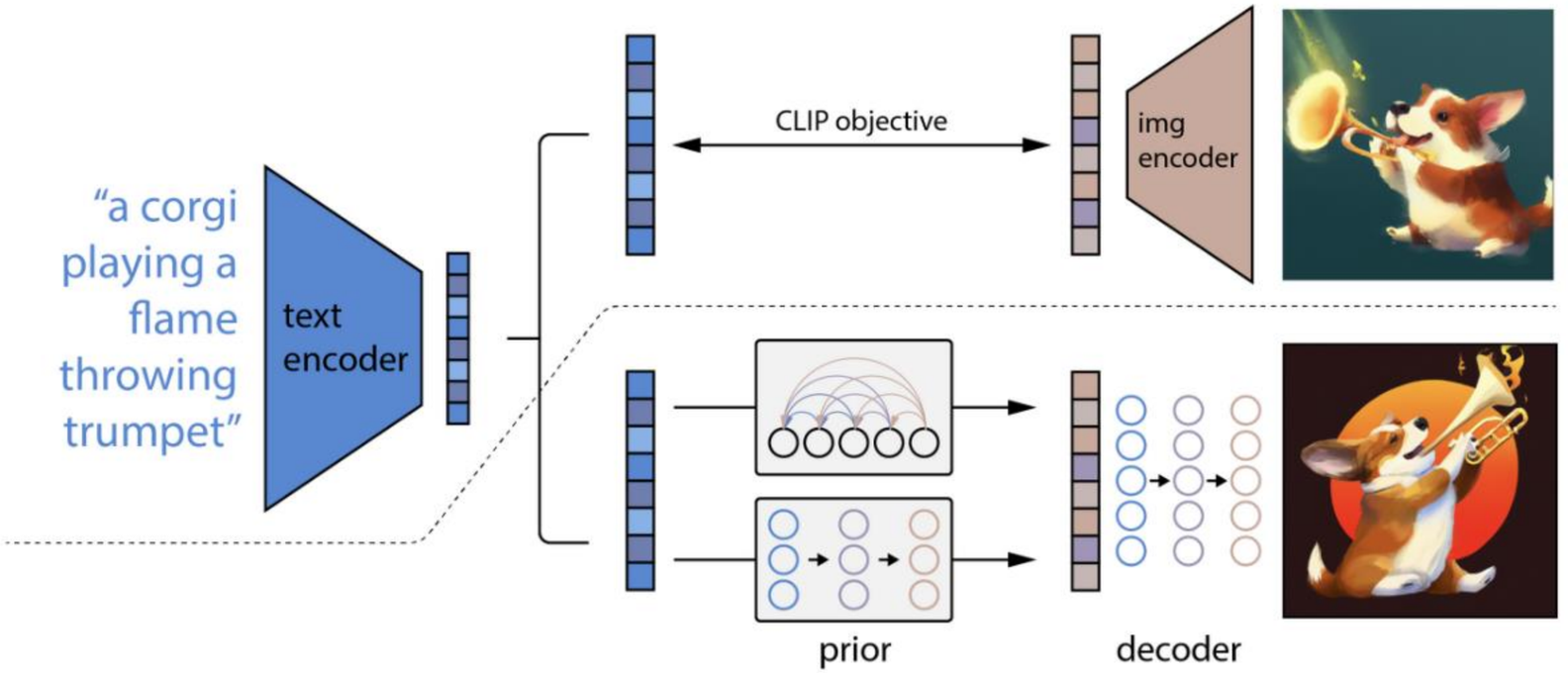
Diffusion Models Beat GANs on Image Synthesis

Prafulla Dhariwal*
OpenAI
prafulla@openai.com

Alex Nichol*
OpenAI
alex@openai.com

03 DALL-E의 기술- unCLIP

unCLIP



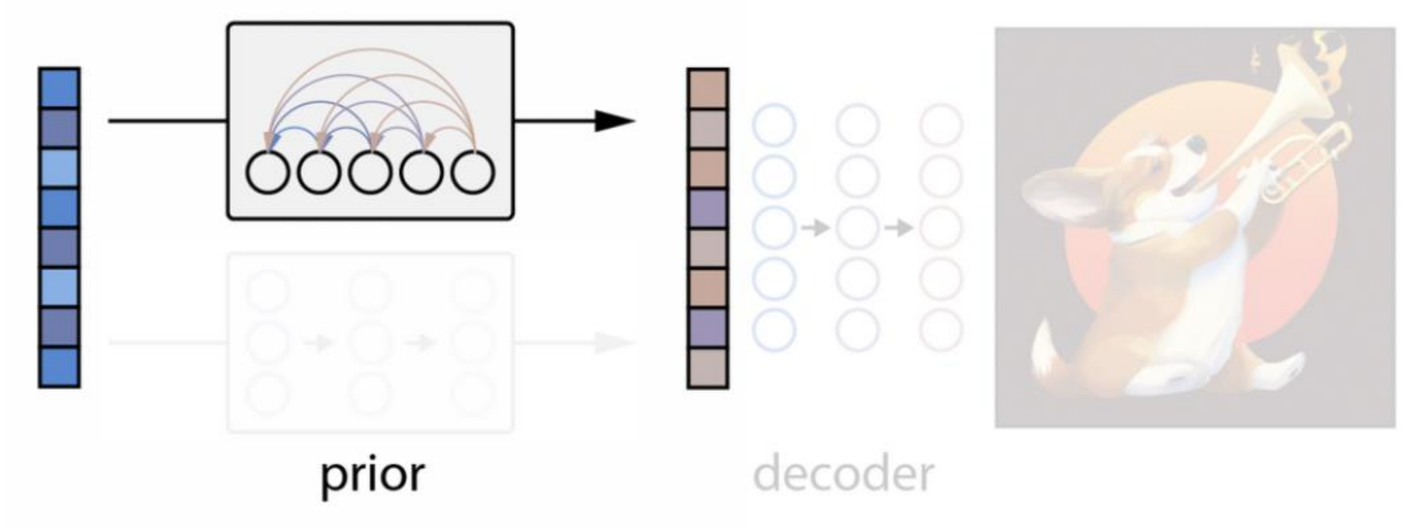
1. CLIP pretraining

2. DIFFUSION

prior & decoder

03 DALL-E의 기술- unCLIP

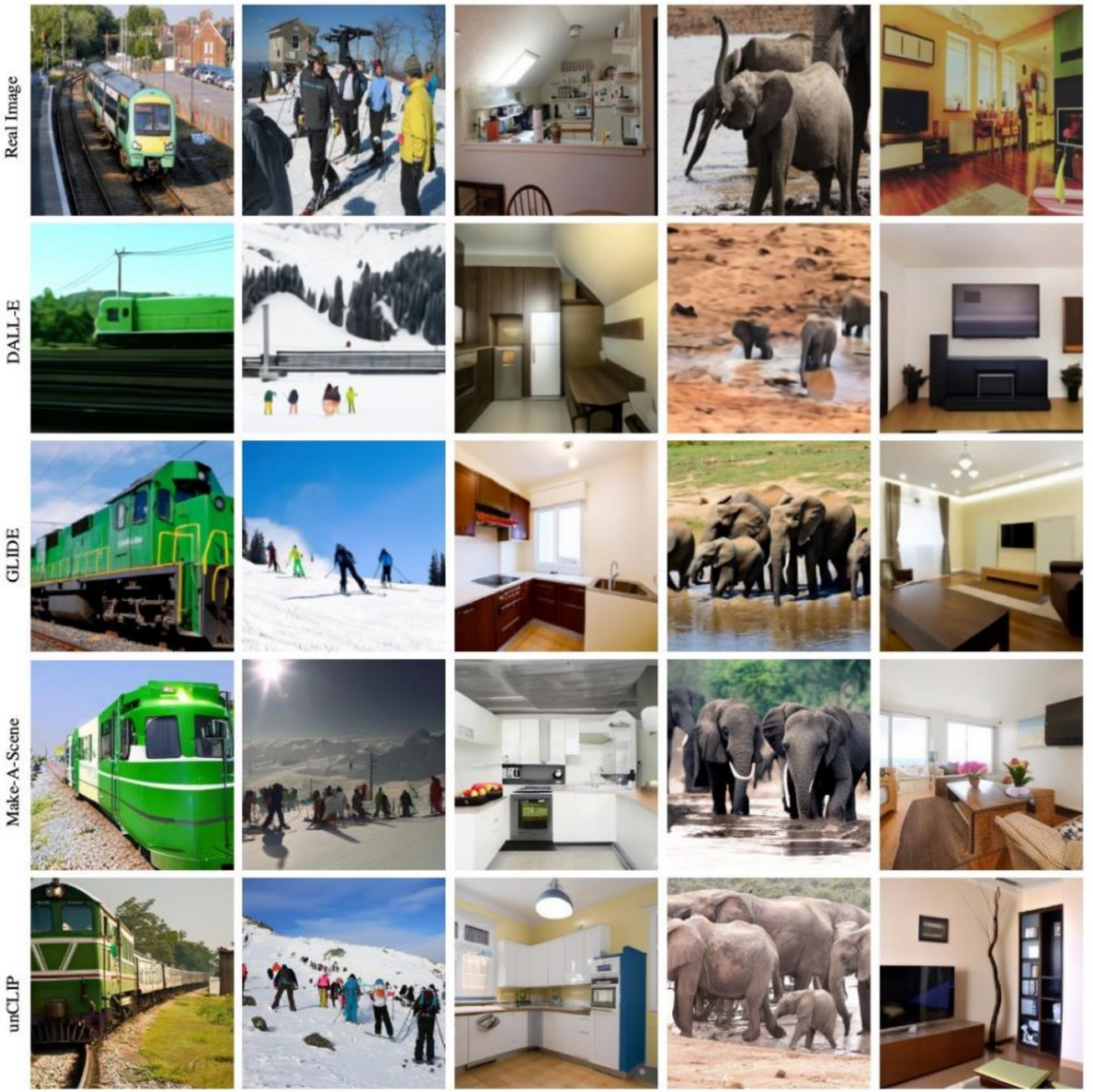
prior



03 DALL-E의 기술- unCLIP

prior

03 DALL-E 2


Real Image
DALL-E
GLIDE
Make-A-Scene
unCLIP
"a green train is coming down the tracks"
"a group of skiers are preparing to ski down a mountain."
"a small kitchen with a low ceiling"
"a group of elephants walking in muddy water."
"a living area with a television and a table"

03 DALL-E의 기술- 주식 시장