

# 챗GPT 완벽정리 - 영상처리

2023. 10.

정 준 수 PhD

# 생성형AI는 어떻게 이미지를 만들까요?

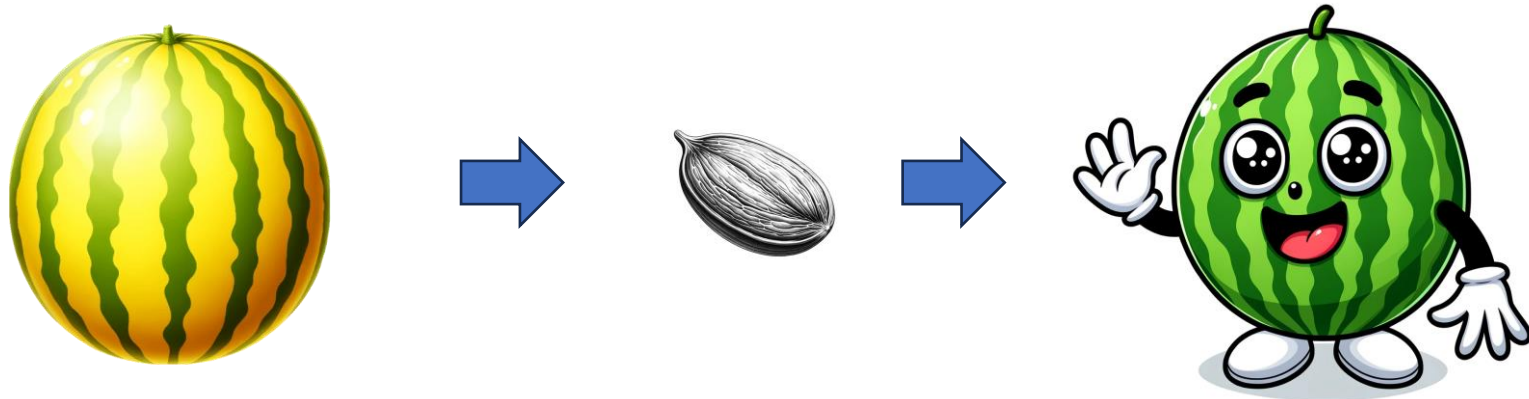
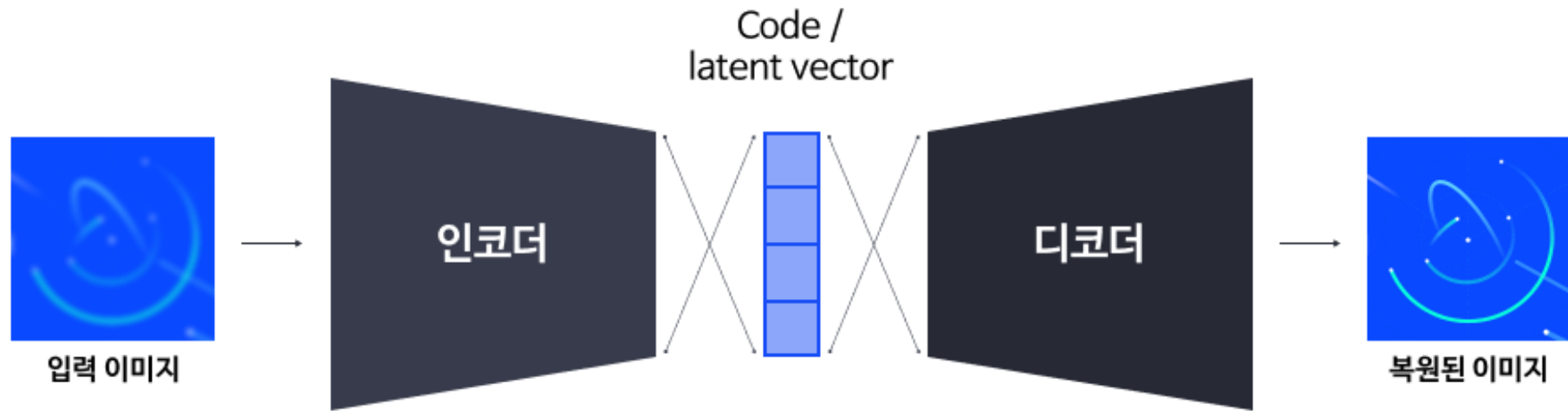
## 1. 인식모델과 생성모델의 차이 비교

생성형 AI 기술은 인공신경망을 활용하여 다양한 데이터를 학습하고, 이를 바탕으로 사람처럼 문맥과 의미를 파악하여 새로운 데이터를 자동으로 생성하는 기술이다. 기존의 AI 기술은 주로 데이터의 판별을 목적으로 회귀, 분류, 군집화와 같은 작업을 수행했으나, 생성형 AI 기술은 사용자의 질문이나 과제 해결을 위해 데이터에서 패턴과 규칙을 학습하며 새로운 콘텐츠를 생성한다(국가정보원 외, 2023).

## 인식모델과 생성형 모델 비교

비교 기준	인식 모델 (Recognition Model)	생성 모델 (Generative Model)
주된 목적	데이터 분류 및 군집화	새로운 데이터 생성
데이터 변환	고차원 원본 데이터를 저차원 라벨 데이터로 변환	사람과 유사한 방식으로 문맥과 의미를 이해하고 새로운 데이터 생성
예시	음성 인식(Speech To Text), 문자 인식(Scene Text Recognition)	이미지, 비디오, 오디오, 텍스트 생성
학습의 기반	주어진 데이터의 패턴과 규칙 학습	대량의 데이터 학습을 통해 패턴과 규칙 학습
사용하는 기술	인공신경망, 딥 러닝	인공신경망, 딥 러닝, 생성적 적대 신경망 (GANs)
결과	구조화된 저차원 라벨 데이터	새롭고 다양한 고차원 데이터

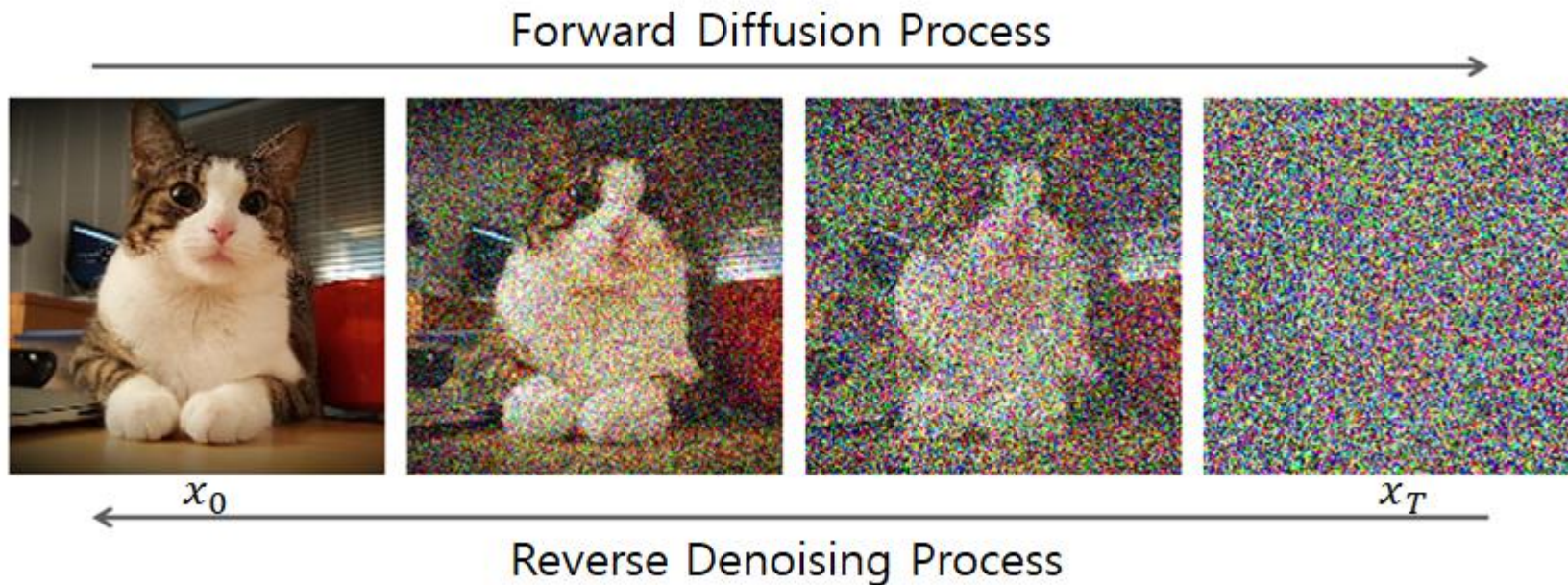
## 인코더와 디코더 모델

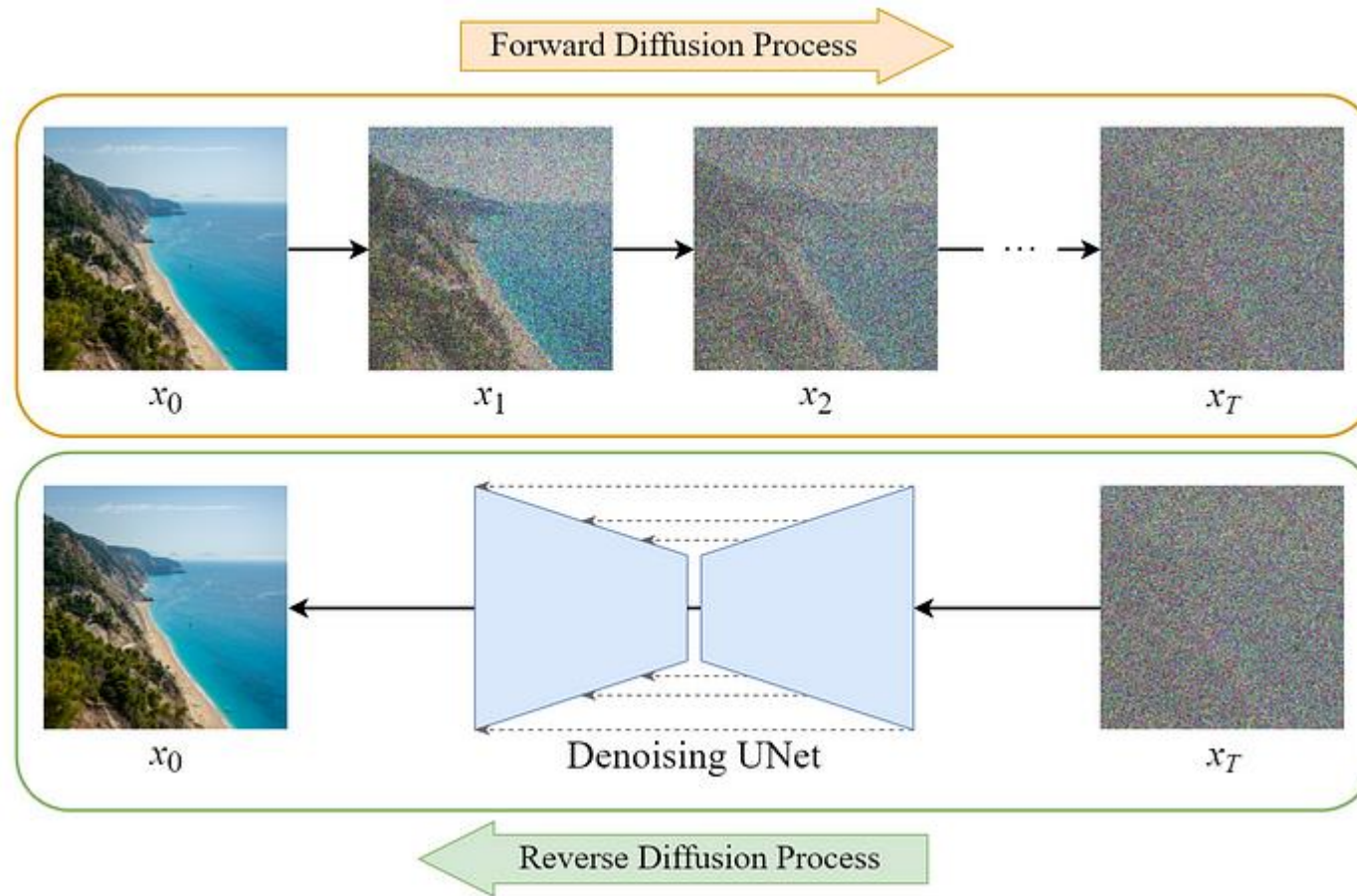


# 생성형AI는 어떻게 이미지를 만들까요?

## 2. Diffuser 모델의 이해

Diffuser 모델은 입력 이미지에 정규 분포의 Noise를 여러 단계로 추가하고, 다시 여러 단계로 Noise를 제거하여, 입력 이미지와 유사한 확률 분포의 결과 이미지를 생성합니다.





이미지에서 Noise가 점차 확산되는 현상 때문에 이를 'Diffusion'이라 부릅니다.

Forward Diffusion Process에서는 고정된 Gaussian 분포로 생성된 Noise를 이미지에 추가하며, Reverse Diffusion Process에서는 학습된 Gaussian 분포로 생성된 Noise 이미지를 원본 이미지에서 제거합니다.

**GPT는 왜 같은 Prompt에 다른 답을 줄까요?**



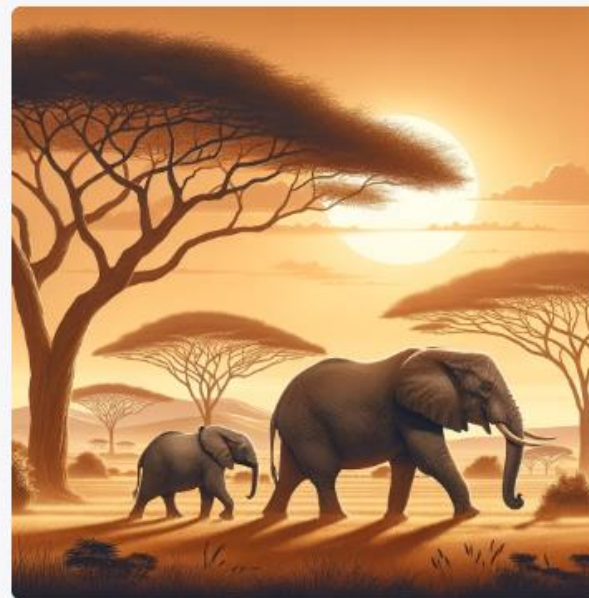
JS

아프리카 초원을 걷는 2 마리 코끼리를 그려 주세요



DALL-E 3

Created 2 images



여기 아프리카 초원을 걷는 두 마리 코끼리의 이미지를 제공합니다. 어떻게 생각하시나요?





DALL-E 3  
Created image



죄송합니다. 요청하신 내용 중 일부가 저희의 "content policy"에 따라 이미지를 생성할 수 없었습니다. 다른 요청을 해주시면 도와드리겠습니다.

JS

에니메이션 풍으로 그려주세요



DALL-E 3

Created 2 images



에니메이션 스타일로 그려진 아프리카 초원 위의 코끼리 이미지를 보여드립니다. 만족하시나요?

JS

에니메이션 풍으로 그려주세요



DALL-E 3

Created 2 images



에니메이션 스타일로 그려진 아프리카 초원의 코끼리 이미지를 제공합니다. 어떠신가요?





gorgeous cute Austrian girl, **smiling**, (crop top), Steel gray hair loose braided hair, comics printed shirt, oversized jogger pants, lean against the wall, backstreet



gorgeous cute Austrian girl, **angry**, (crop top), Steel gray hair loose braided hair, comics printed shirt, oversized jogger pants, lean against the wall, backstreet



**backstreet**



**meadows**





**laughing**



**angry**

**원하는 결과를 얻기 위한 방법들?**





face focus, cute, 1girl, black background,  
light particle, solo, standing, cinematic  
composition, best lighting, looking up



smiling



**angry**



**sad**





**laughing**



**smiling(random seed)**

# Inpainting





# 같은 얼굴각도의 영상을 얻기 위한 Canny ControlNet



input image



Canny image



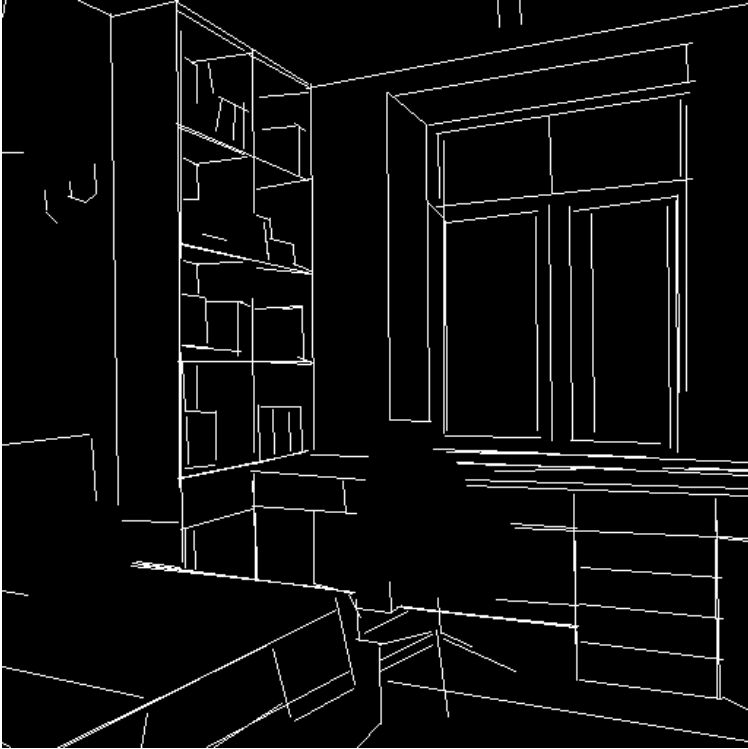
Result 0



Result 1

**Girl with Pearl Earring**  
**(Generated images with original Stable Diffusion v1.5 + canny ControlNet)**

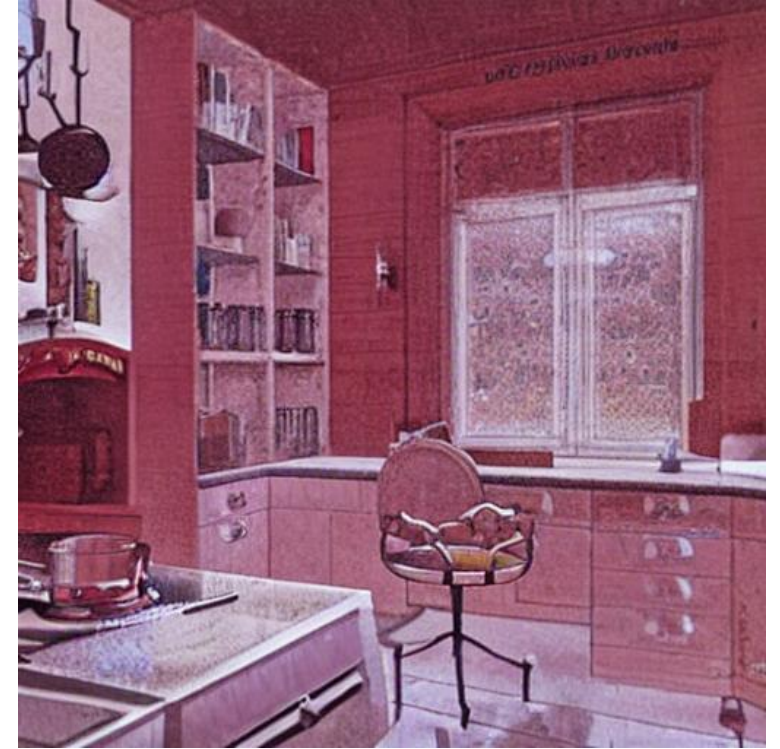
# 원하는 구조의 영상을 얻기 위한 controlnet-mlsd



Template



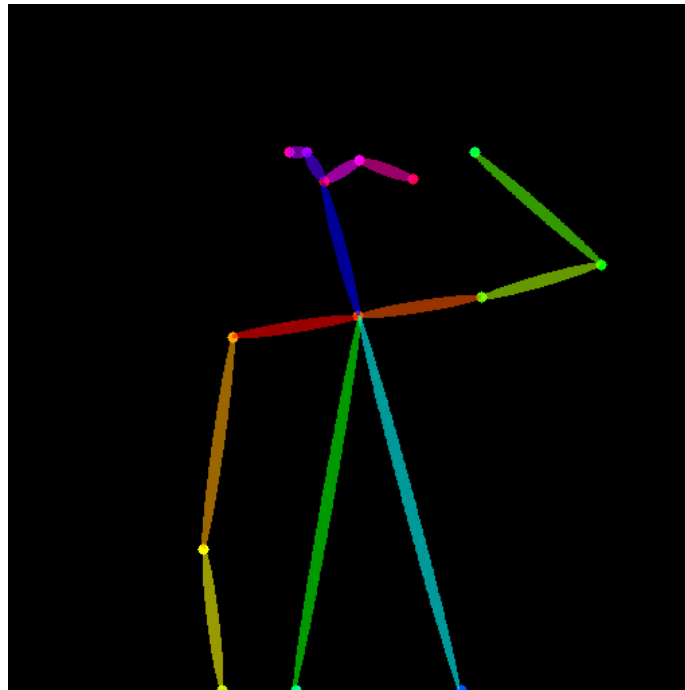
Modern office



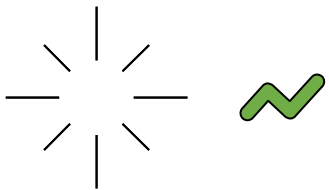
Classical kitchen



# 원하는 자세 영상을 얻기 위한 controlnet-openpose







## 강사 소개

정 준 수 PhD

jsjeong@hansung.ac.kr

Thank you!

### 【학력】

- 고려대학교 전기공학사
- 뉴욕 공대 전산학석사(AI 전공)
- 한성대학교 컨설팅학박사

### 【경력】

- (前) 삼성전자, 삼성의료원, 삼성SDS 연구원
- (前) 한성대학교 겸임교수 – 머신러닝과 인지과학 강의
- (現) (주)퍼즐시스템즈 연구소장

