

CUAI 8기 논문 발표

# MAISI: Medical AI for Synthetic Imaging

2022.03.25

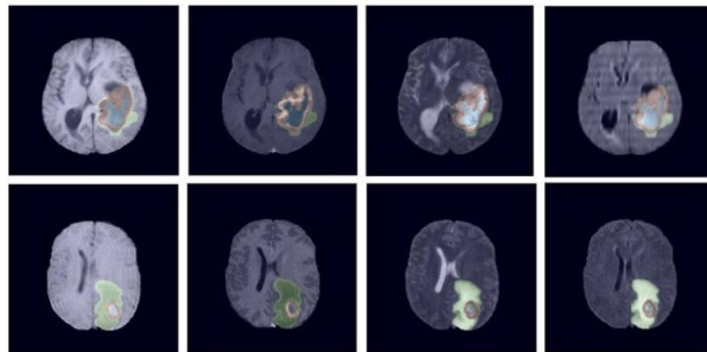
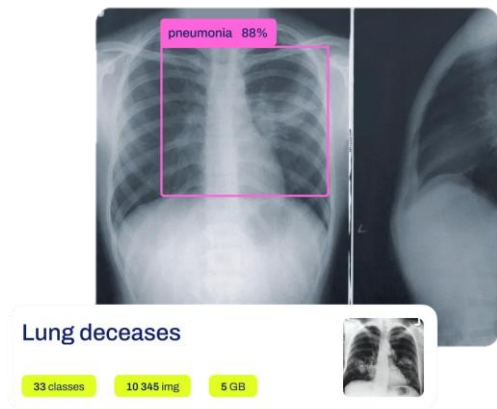
발표자 : 정현석

## 목차

1. Medical AI 이란
2. Medical AI's Challenges
3. Synthetic Data in Medical AI
4. 논문 이해를 위한 3가지 사전 지식
5. 논문 소개 (MAISI: Medical AI for Synthetic Imaging)
6. MAISI's 3 development stages
7. Evaluation
8. Data Augmentation in Downstream Tasks

## Medical AI

Medical AI는 환자의 CT나 MRI 같은 의료 이미지를 분석해서 질병을 진단하거나 치료 계획을 세우는 데 쓰입니다.



# Medical AI's Challenges

## 1. Data scarcity

## 2. High annotation cost

## 3. Privacy concern



AI타임스

<https://www.aitimes.com> , news , articleView ;

### 의료 AI 데이터 편향 심각하다...미국과 중국 데이터가 절반 ...

2022. 4. 9. — 임상 AI에 사용되는 데이터 세트의 절반 이상이 미국이나 중국에서 온 것이라는 연구 결과가 나왔다. AI가 알고리즘을 훈련하고 검증한 결과는 데이터 ...



메디팜스투데이

<https://www.pharmstoday.com> , 뉴스 , 지구촌통신 ;

### FDA 승인 AI 기반 의료기기 40% 이상 검증 데이터 부족

2024. 9. 10. — 2022년까지 미국 FDA에서 승인된 모든 AI(인공지능) 기반 기기의 약 43%가 검증 데이터가 부족하고, 28%만이 전향적 테스트를 받았다는 연구결과가 ...



tradingclue.kr

<https://tradingclue.kr> , medical-ai-bias ;

### 의료 AI의 편향

2024. 7. 10. — 최근 연구에 따르면, 의료 AI 모델은 특정 인구 집단에서 일관되게 낮은 성능을 보이는 경향이 있으며, 특히 여성과 유색인종에서 이러한 문제가 두드러 ...



한겨레

<https://www.hani.co.kr> , economy , economy\_general ;

### '내 CT사진' AI가 맘대로 쓴다고?...의료 민감정보 활용, 물꼬 ...

2024. 2. 4. — 가이드라인에 따르면, 병원 기업 등 개인정보 처리자는 시트, 엑스레이, 엠알아이(MRI) 같은 의료데이터에 담긴 환자번호·생년월일·성별 등 환자 관련 ...



메디칼타임즈

<https://www.medicaltimes.com> , News , NewsView ;

### 의료 AI 도입만으로 뇌경색 예후 월등...한계는 비용 문제

2022. 12. 20. — 국내 다기관 연구진, 'RAPID AI' 도입 전후 비교 분석 연구 예후 좋은 환자 비율 10%→62.5% 경증..."도입 기관 늘려야".

## Synthetic Data in Medical AI

그래서 실제처럼 보이는 가짜 데이터,  
즉 합성 데이터(Synthetic Data)가 필요합니다.

## Synthetic Data in Medical AI

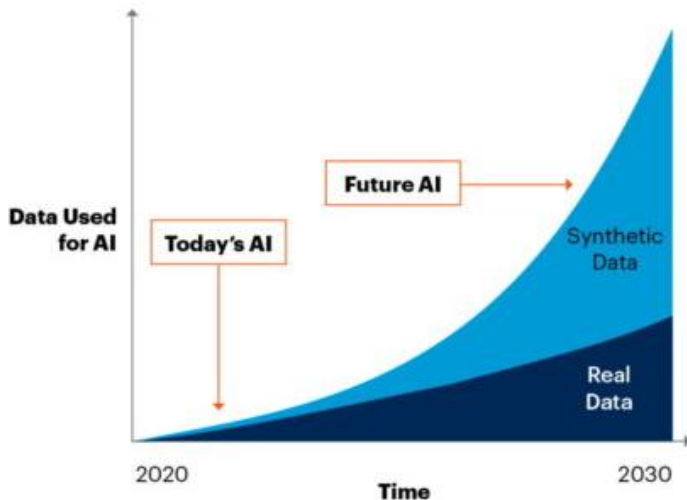
그래서 실제처럼 보이는 가짜 데이터, 즉 합성 데이터(Synthetic Data)가 필요합니다.

### What Is Synthetic Data?

Synthetic data is annotated information that computer simulations or algorithms generate as an alternative to real-world data.

Put another way, synthetic data is created in digital worlds rather than collected from or measured in the real world.

It may be artificial, but synthetic data reflects real-world data, mathematically or statistically. Research demonstrates it can be as good or even better for training AI model than data based on actual objects, events or people.



## 논문 소개

### MAISI: Medical AI for Synthetic Imaging

Pengfei Guo\*  
NVIDIA

Can Zhao\*  
NVIDIA

Dong Yang\*  
NVIDIA

Ziyue Xu  
NVIDIA

Vishwesh Nath  
NVIDIA

Yucheng Tang  
NVIDIA

Benjamin Simon  
National Institutes of Health  
University of Oxford

Mason Belue  
University of Arkansas for Medical Sciences

Stephanie Harmon  
National Institutes of Health

Baris Turkbey  
National Institutes of Health

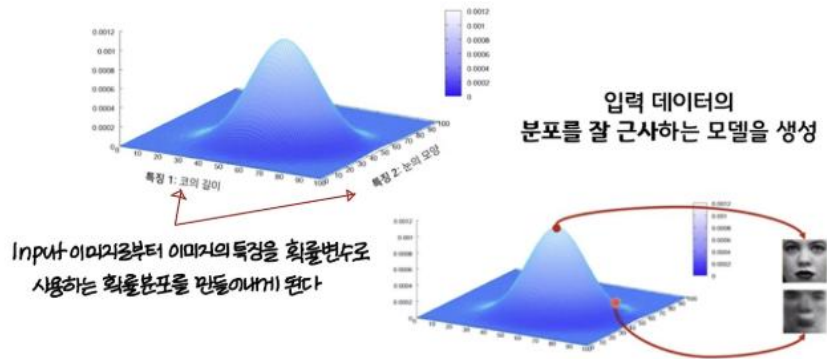
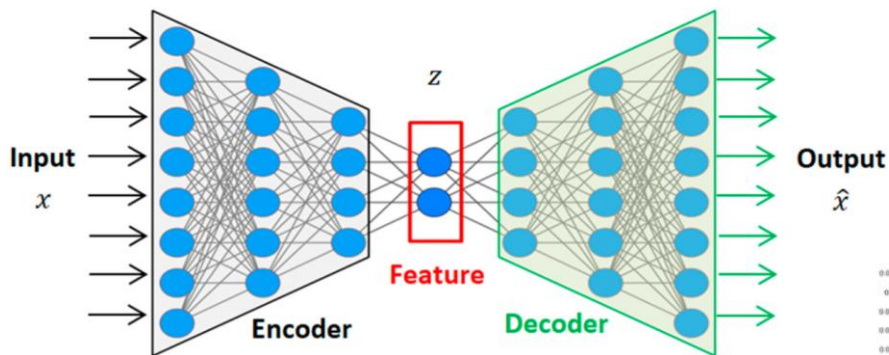
Daguang Xu  
NVIDIA



*Pengfei Guo, Can Zhao, Dong Yang, Ziyue Xu, Vishwesh Nath, Yucheng Tang, Benjamin Simon, Mason Belue, Stephanie Harmon, Baris Turkbey, Daguang Xu*; Proceedings of the Winter Conference on Applications of Computer Vision (WACV), 2025, pp. 4430-4441

## VAE, Variational Auto-Encoder (사전지식 1/3)

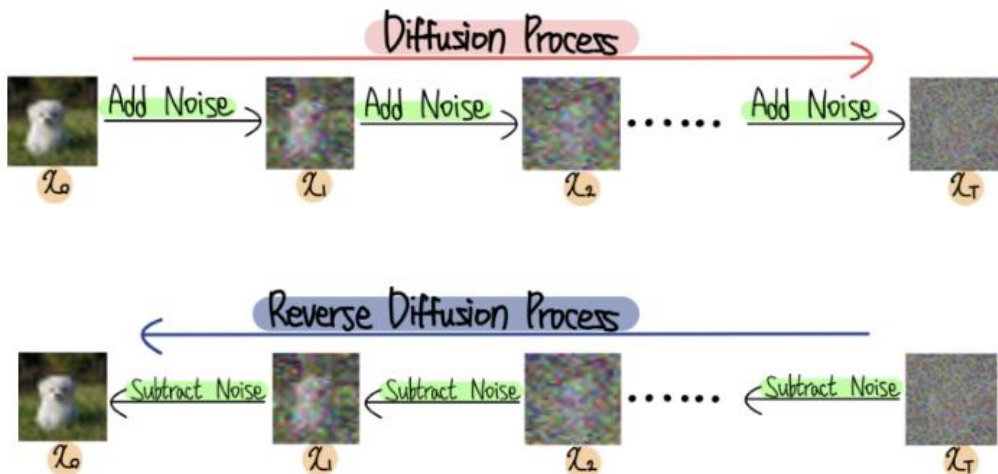
- Variational Auto-Encoder(VAE)는 생성 모델
- VAE는 Input image  $X$ 를 잘 설명하는 feature를 추출하여 Latent space에 담고, 이를 통해 새로운 데이터를 생성



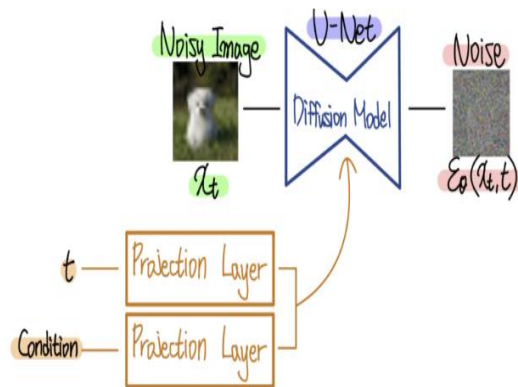


## Diffusion Model (사전지식 2/3)

- Diffusion Model도 생성 모델!
- '우리가 만약 어떤 데이터에 랜덤한 노이즈를 조금씩 조금씩 뿌려서 완전한 노이즈를 만들 수 있다면, 거꾸로 완전한 노이즈에서 처음 데이터를 복원하는 것도 가능하지 않을까?'



Diffusion Process 설명



Diffusion Model Architecture

## ControlNet (사전지식 3/3)

- **ControlNet은 Diffusion 모델의 생성 과정을 '내가 원하는 방식'으로 제어할 수 있게 해줌**

### 1. Stable Diffusion으로 생성한 "테니스 치는 아이언맨"



## ControlNet (사전지식 3/3)

- ControlNet은 Diffusion 모델의 생성 과정을 '내가 원하는 방식'으로 제어할 수 있게 해줌

### 1. Stable Diffusion으로 생성한 "테니스 치는 아이언맨"



### 2. 근데 내가 "포핸드 자세로 스윙하는 아이언맨" 를 생성하고 싶다면?

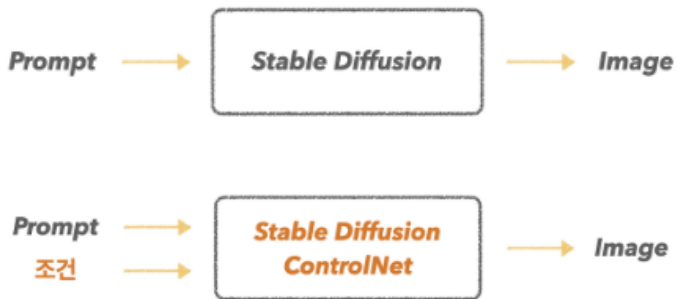


포핸드 자세가 한두개도 아니고, 우리가 원하는 자세를 prompt만으로 제어하는건 어렵다ㅠㅠ

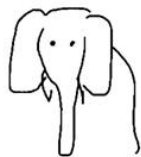
## ControlNet (사전지식 3/3)

Stable Diffusion은 텍스트(prompt)만으로 이미지를 생성했는데,

**ControlNet은 여기에 추가로 이미지 조건 (e.g. Edge, Pose, Depth 등) 함께 넣어서 → 더 정밀한 제어를 가능**



Input (User Scribble)



Default



User Prompt



"magic hot air balloon over a lit magic city at night"



"Egyptian elephant sculpture"

## MAISI: Medical AI for Synthetic Imaging

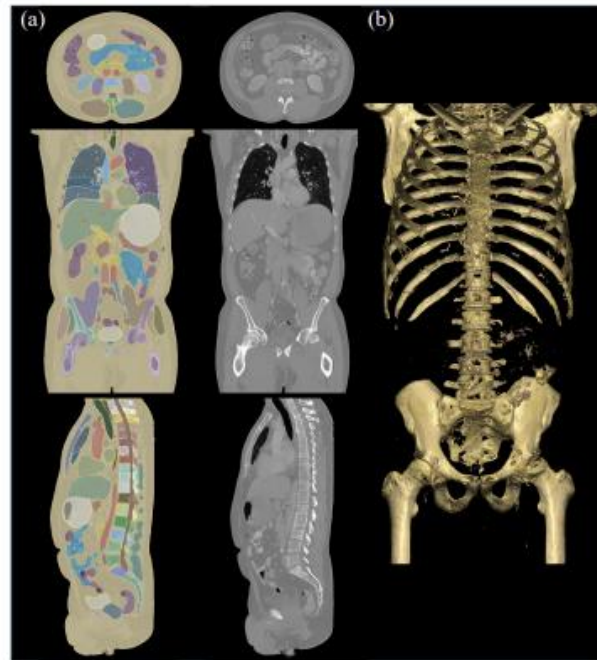
**자! 이제 모든 준비가 끝났습니다.**

**이제 앞에서 배운 3가지를 합치면 이번 논문이 거의 완성됩니다!**

# MAISI: Medical AI for Synthetic Imaging

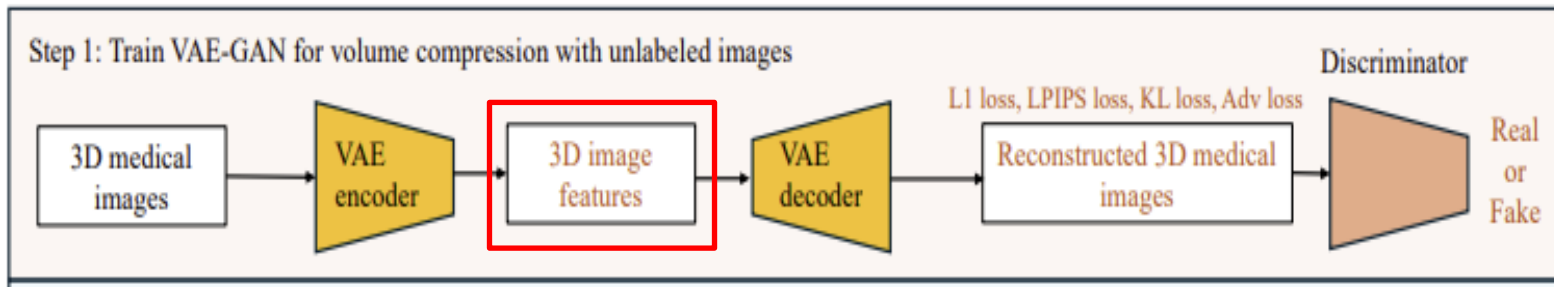
## Abstract

*Medical imaging analysis faces challenges such as data scarcity, high annotation costs, and privacy concerns. This paper introduces the Medical AI for Synthetic Imaging (MAISI), an innovative approach using the diffusion model to generate synthetic 3D computed tomography (CT) images to address those challenges. MAISI leverages the foundation volume compression network and the latent diffusion model to produce high-resolution CT images (up to a landmark volume dimension of  $512 \times 512 \times 768$ ) with flexible volume dimensions and voxel spacing. By incorporating ControlNet, MAISI can process organ segmentation, including 127 anatomical structures, as additional conditions and enables the generation of accurately annotated synthetic images that can be used for various downstream tasks. Our experiment results show that MAISI's capabilities in generating realistic, anatomically accurate images for diverse regions and conditions reveal its promising potential to mitigate challenges using synthetic data.*



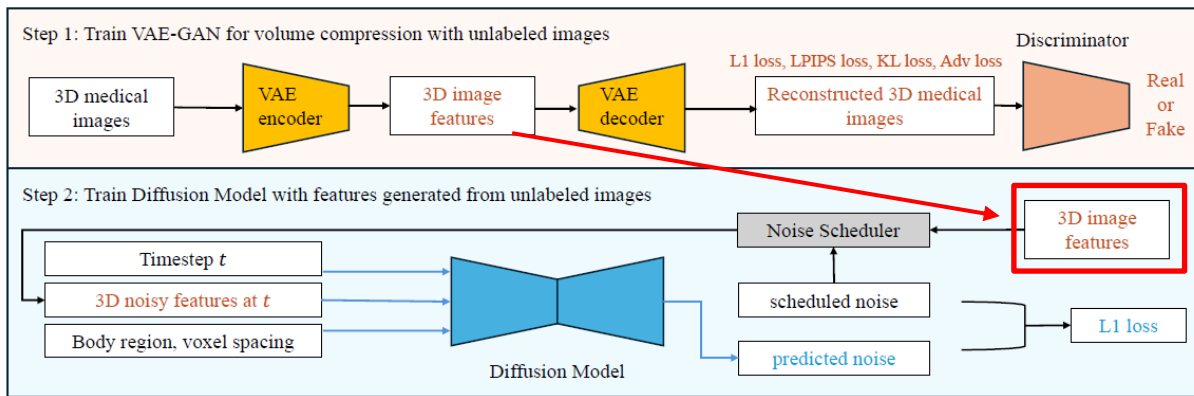
**VAE 기반 압축과 Diffusion 모델을 결합하여, 고해상도 3D CT 합성 이미지를 만들어냅니다.**

## Step 1: VAE을 통해 3D CT를 압축해도 의미 있는 표현을 유지하면서 복원할 수 있도록 학습



1. 3D CT 이미지를 VAE Encoder를 통해 Latent space 으로 압축 (= 3D image features)
2. Decoder가 다시 이를 복원하여 원본 이미지와 최대한 유사하게 생성
3. 복원된 이미지는 Discriminator(판별기)를 통해 진짜/가짜 여부 평가

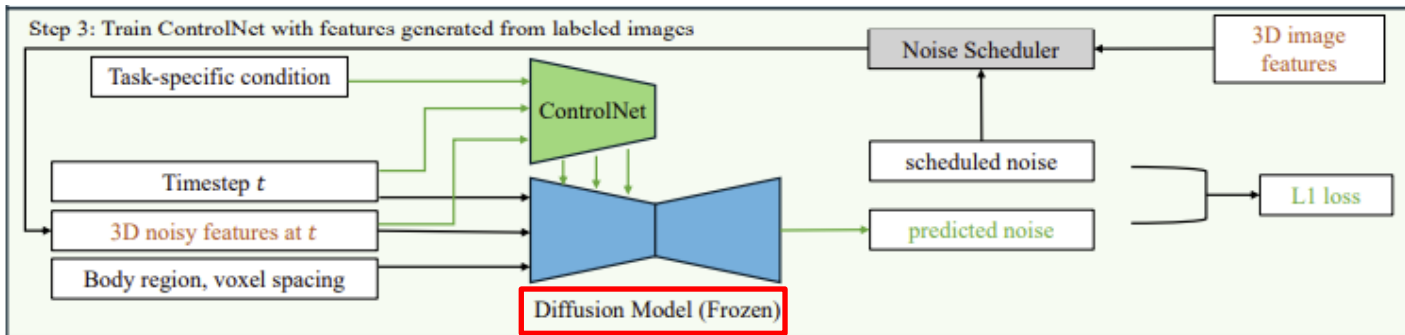
## Step 2: 압축된 latent feature 위에서 Diffusion 모델이 노이즈를 제거하면서 이미지를 복원하는 과정 학습



1. Step1에서 추출된 Latent features에 인위적으로 노이즈 추가
2. 모델은 이 Noise featur와 각 timestep과 해부학적 조건(body region, voxel spacing) 입력 받음
3. 이 과정을 통해 Diffusion 모델은 CT를 복원할 수 있는 Reverse diffusion 과정을 학습한다.

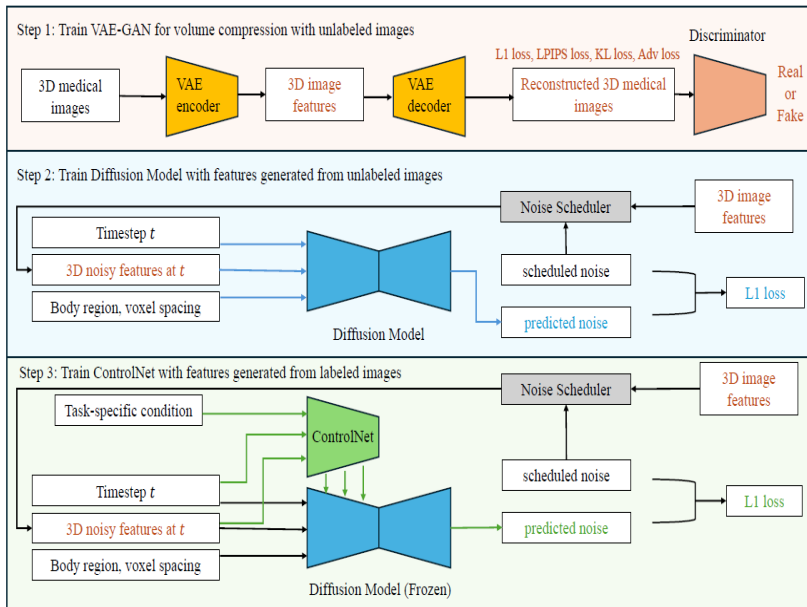


### Step 3 : ControlNet을 통해 조건에 따라 이미지를 정밀하게 생성



1. 이미 학습된 Diffusion모델을 Freeze한 상태에서 조건(e.g. segmentation mask)을 ControlNet에 입력
2. ControlNet의 조건과 Latent Feature를 사용해서 Diffusion 모델이 이미지를 생성한다.
3. 이렇게 해부학적 조건에 맞춘 제어 가능한 CT 이미지 생성이 가능해진다~!

# MAISI's 3 development stages



**Step 별 핵심 Point!**

**Step1: 3D Image Feature 압축**

**Step2: Diffusion 모델 학습만 진행 (생성 X)**

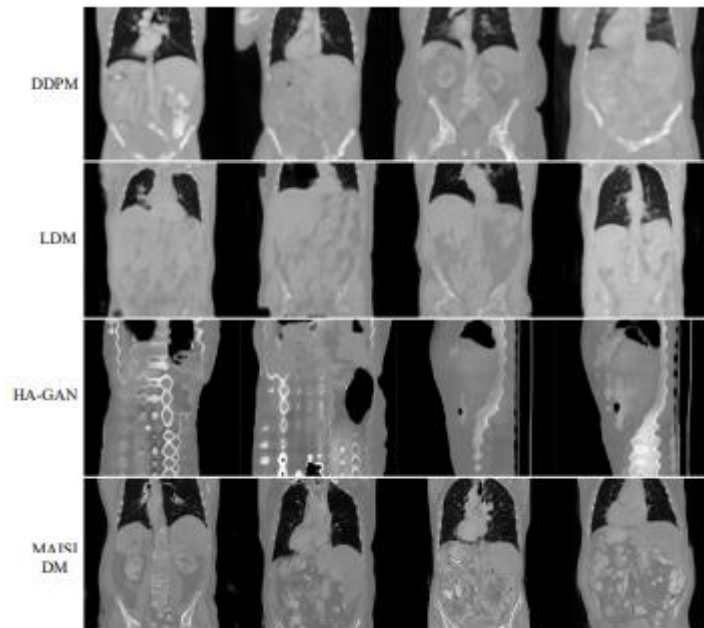
**Step3: ControlNet으로 우리가 원하는 이미지 생성!**

# Evaluation

## Evaluation of MAISI Diffusion Model

Method	FID ↓ (Axial)	FID ↓ (Sagittal)	FID ↓ (Coronal)	FID ↓ (Avg.)
DDPM [29]	18.524	23.696	25.604	22.608
LDM [49]	16.853	10.191	10.093	12.379
HA-GAN [61]	17.432	10.266	13.572	13.757
MAISI DM	<b>3.301</b>	<b>5.838</b>	<b>9.109</b>	<b>6.083</b>

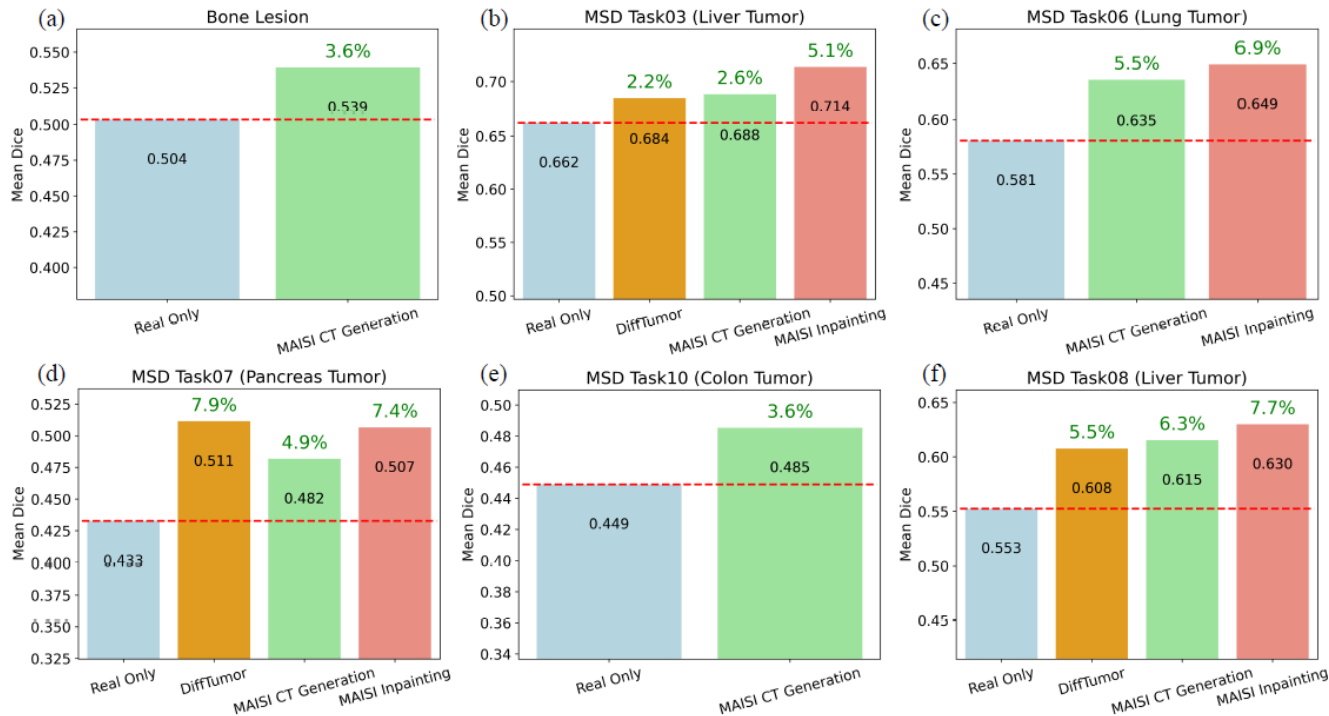
Table 3. Fréchet Inception Distance across three views between MAISI DM and retrained baseline methods using the unseen dataset autoPET 2023 [20] as the reference.



\*FID (Fréchet Inception Distance): 생성된 이미지가 실제 이미지와 얼마나 유사한지를 수치로 평가하는 지표

# Data Augmentation in Downstream Tasks

생성된 Synthetic CT Data를 추가해 학습했을 때, segmentation 성능이 유의미하게 향상



단순히 "잘 만들어진 이미지"가 아니라, → 실제로 Medical AI 모델의 성능 향상에 기여한다!

구경하고 가세요~

