

# 강화학습 과제

A69025 박석래

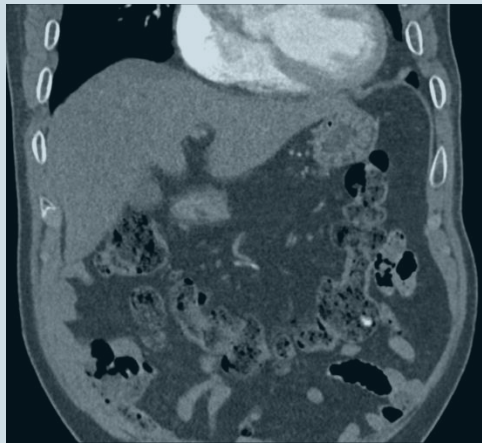
# 개요

1. 프로젝트 주제 및 목표
2. 환경 및 데이터셋
3. State, Action, Reward
4. 강화학습 알고리즘 및 hyperparameter
5. 실험 및 결과 (test01~07)
6. 토의 및 결론

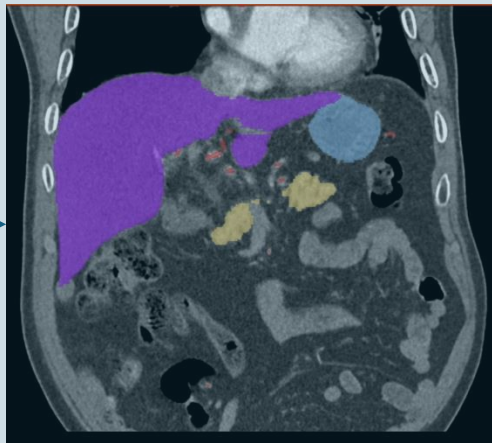
# 1. 프로젝트 주제 및 목표 - 강화학습을 활용한 의료영상 Segmentation

- 회사에서 서비스 중인 제품의 제작 공정 효율화

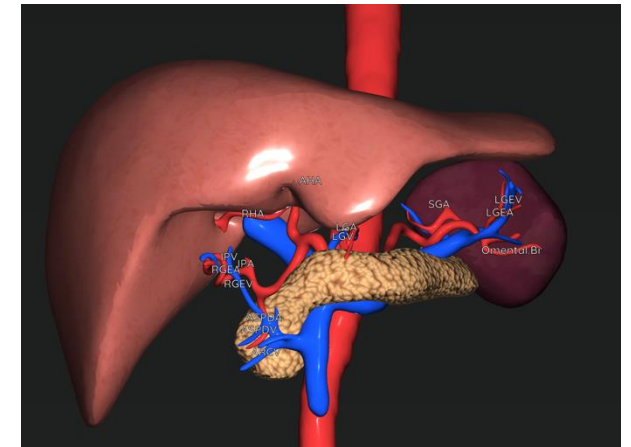
## 프로젝트 타겟 제작 과정



수술 예정 환자의 CT DICOM



해당 수술의 필수 구조물  
segmentation

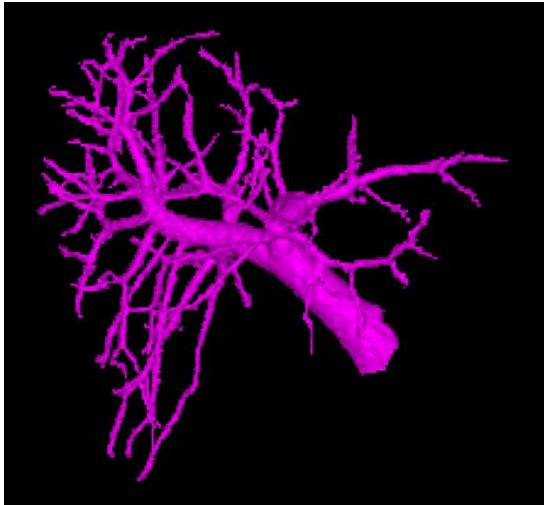


수술 전 계획 및 내비게이션 제품

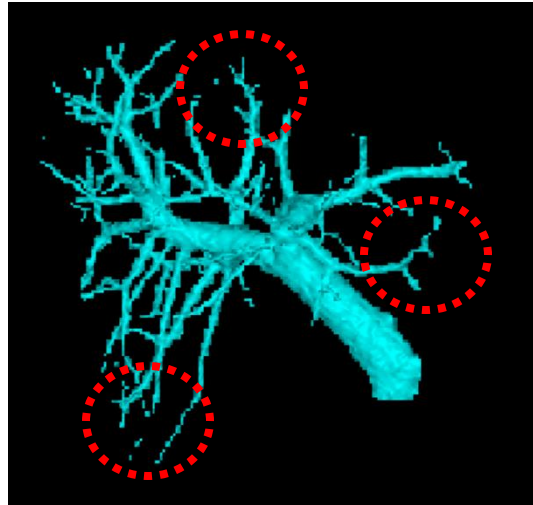
## 전체 제작 과정

# 1. 프로젝트 주제 및 목표

- 현행 문제점 및 개선 방안



Ground truth



nn-Unet 기반의  
segmentation 알고리즘

혈관 구조에서,  
말단부 혈관으로 갈수록 연결성에  
대한 정확도 부족으로 인한 제작  
시간 증가

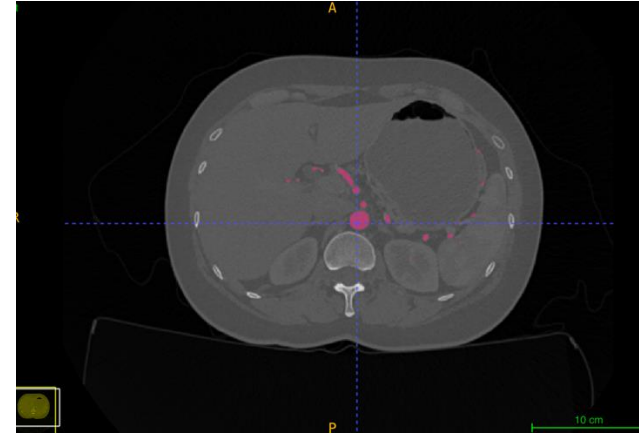


혈관 구조에서,  
혈관의 연결성을 파악하여,  
segmentation 할 수 있는 알고리즘  
개발

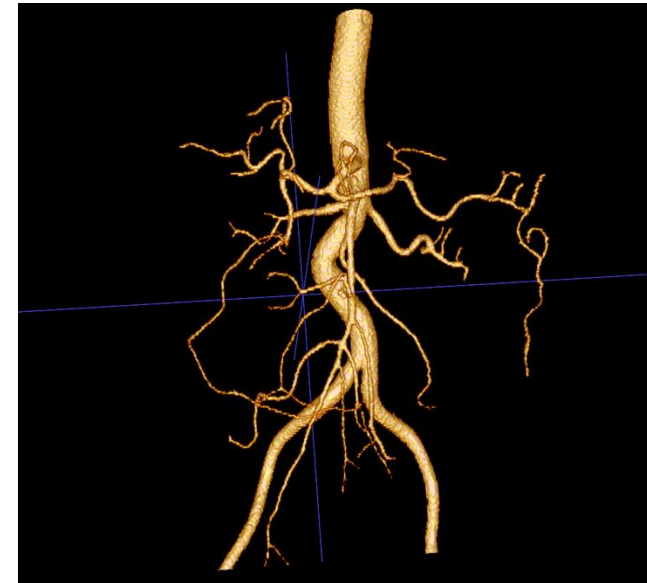
## 2. 환경 및 데이터셋

### A. 데이터셋

- **CT scan (DICOM)** : 99 case
  - Phase : Arterial
  - Range : Abdominal
  - Slice thickness 다양
- **Mask (nifti)** : 99 case
  - Whole artery
    - 위절제술시 필요한 동맥들이 표현되어 있음



한 명의 환자에서 CT scan 과 Mask 의 overlay 예시



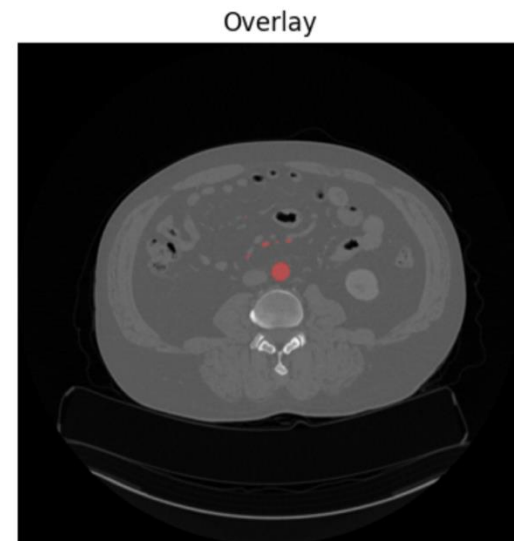
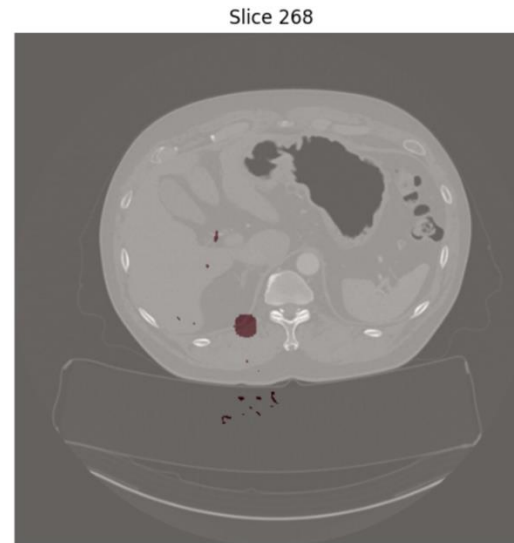
[Whole\_artery.nii.gz] 의 Volume rendering 예시

## 2. 환경 및 데이터셋

### B. 데이터셋 전처리 (CT scan)

- **DICOM to nifti 변환**
  - 환자별 CT scan (DICOM) 을 정렬하여 3D volume 재구성
  - 의료영상 공통 포맷인 **NIfTI (.nii/.nii.gz)** 형태로 저장
    - 변환 시 다음의 값을 유지 (Voxel spacing / Orientation / Slice thickness)
- **Nifti from CT scan 과 Mask 의 공간적 일치(Space Consistency) 확인**
  - 다음을 확인함
    - **Voxel spacing** : CT 와 Mask 의 축별 voxel 크기가 동일해야 함
    - **Orientation (affine matrix)** : 좌우 반전 또는 축 뒤집힘 문제 여부 확인
    - **Volume shape** : 두 nifti 파일의 3D shape 이 일치해야 정확한 segmentation 가능
- **CT 정규화**
  - CT 의 HU(Hounsfield Unit) 값은 매우 넓은 범위를 가지므로, 딥러닝 모델의 학습 효율을 위해 다음과 같은 정규화를 수행함
    - HU Clipping
      - 혈관 영역 학습에 널리 쓰이는 [-200, 300] 으로. Clipping 및 [0,1] 범위로 정규화

(예시) Nifti 간 affine 이 맞지 않는 경우



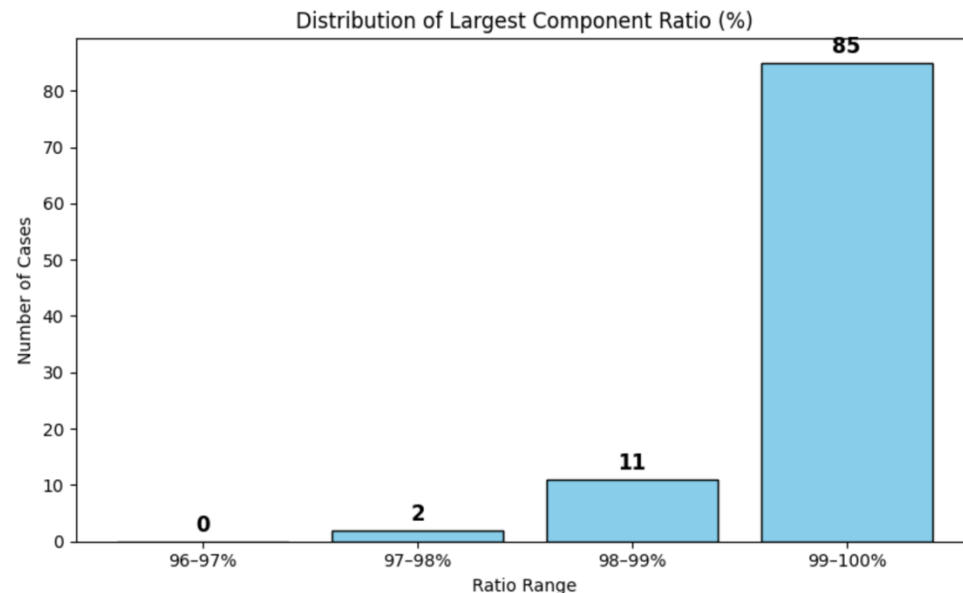
(예시) 전처리 후

## 2. 환경 및 데이터셋

### B. 데이터셋 전처리 (Mask)

- **Mask의 연결성 확인 (Largest component 만 추출)**

- 끊겨있는 혈관은 존재할 수 없음 + 혈관 연결성의 학습이 필요하기 때문에, 끊겨 있는 부분이 있으면 해당 부분은 제외함

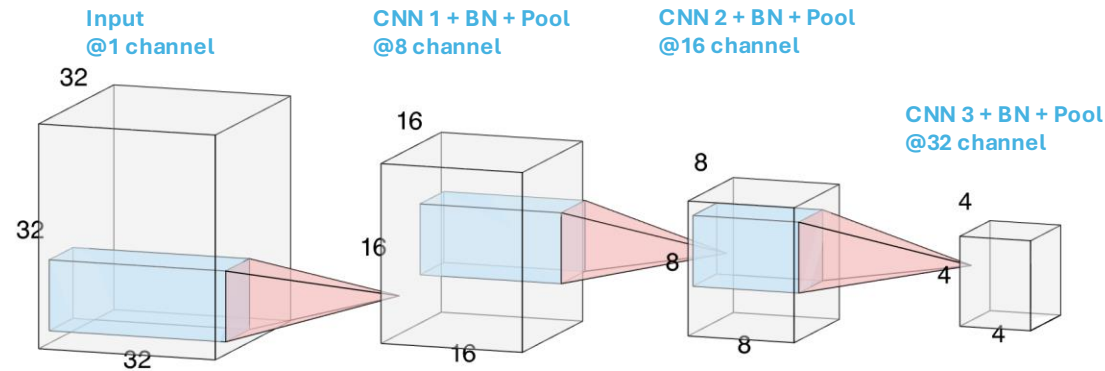


전체 Case	99
Largest component 비율 평균	99.4 %
최소비율	95.8 %
최대비율	99.99%
Component 개수(median)	38 개

### 3. State, Action, Reward 설계

#### A. State (v1)

- Agent의 state는 현재 (시작 시 또는 action 후 이동점) 좌표를 중심으로 한 local 3D patch에서 추출된 CNN feature
- 입력위치
  - (z,y,x) 좌표
- 입력범위
  - 3D patch 크기 (patch size : 32x32x32)



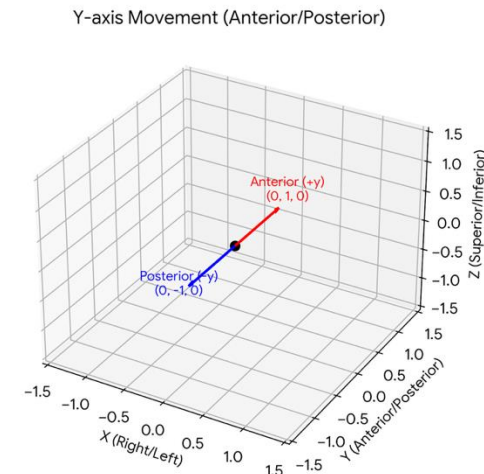
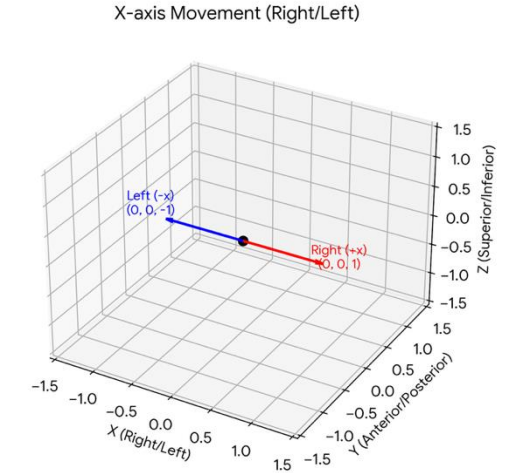
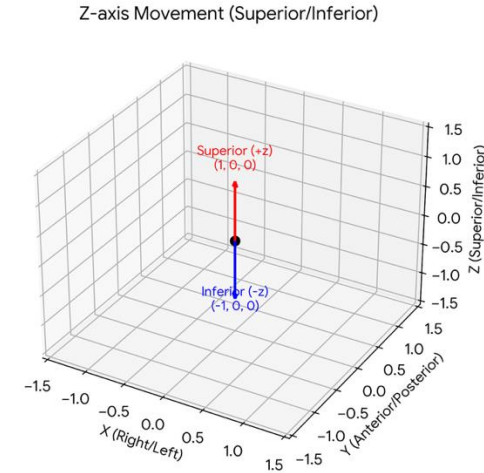
State : CNN encoder



### 3. State, Action, Reward 설계

#### B. Action (v1)

- 6 Discrete action space
  - (z,y,x) 각 방향으로 +1 또는 -1
    - (+1,0,0) : +z, superior
    - (-1,0,0) : -z, inferior
    - (0,+1,0) : +y, anterior
    - (0,-1,0) : -y, posterior
    - (0,0,+1) : +x, right
    - (0,0,-1) : -x, left



### 3. State, Action, Reward 설계

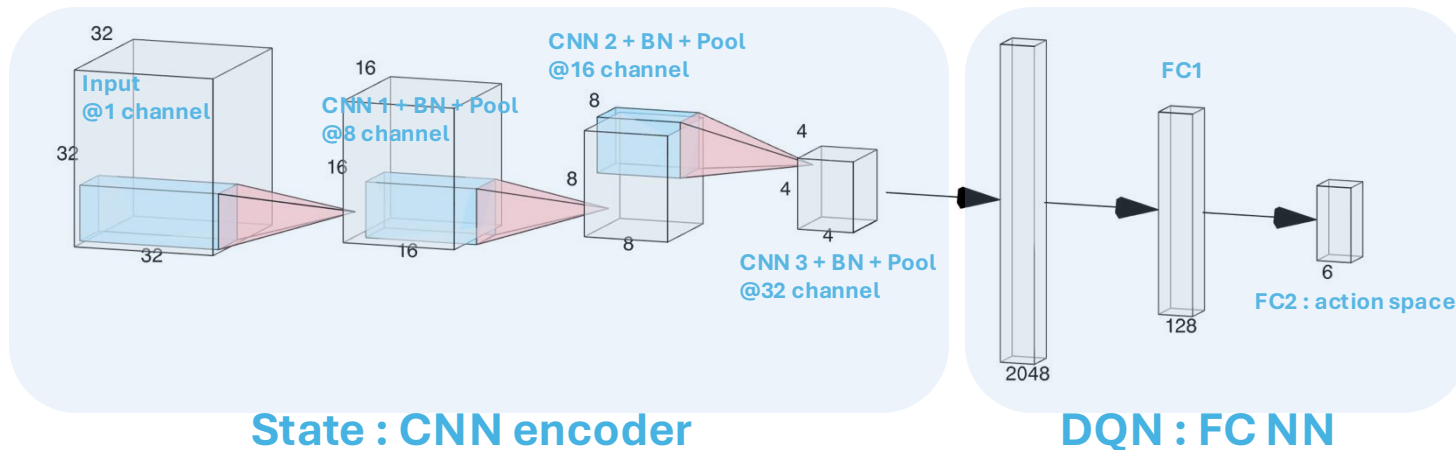
#### C. Reward (v1)

- Agent 가 이동(action) 후 위치(좌표)가 혈관 마스크 내부에 있으면 +1
- 혈관 마스크 외부로 벗어나면 -1 & episode 종료

## 4. 강화학습 알고리즘 구조 (v1)

- **3D-DQN 기반 혈관 추적 (Vessel Tracking)**

- State 입력: 현재 agent 위치 주변의  $32 \times 32 \times 32$  3D patch
- CNN Encoder를 통해 state  $\rightarrow$  2048-D latent vector 로 변환
  - (현재 위치의 CT 패치 3D 인식)
- Latent vector를 DQN에 입력하여 6개 action의 Q-value 계산
- Agent는 가장 큰 Q-value를 갖는 action을 선택하여 다음 위치로 이동



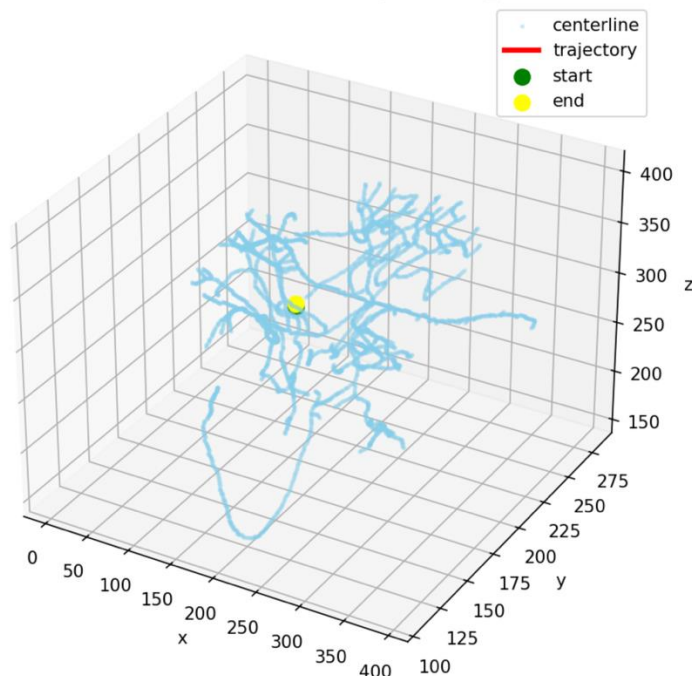
## 5. 실험 (test\_01)

- 모델구조 (v1)
  - State(v1), Action(v1), Reward(v1)
- State 시작 위치 생성기준
  - 혈관 mask 안의 voxel 중 하나 랜덤 선택
- Hyperparameter
  - n\_episodes=200
  - max\_steps=32
  - batch\_size=8
  - Epsilon 선형감소
  - Target network 동기화 : 20

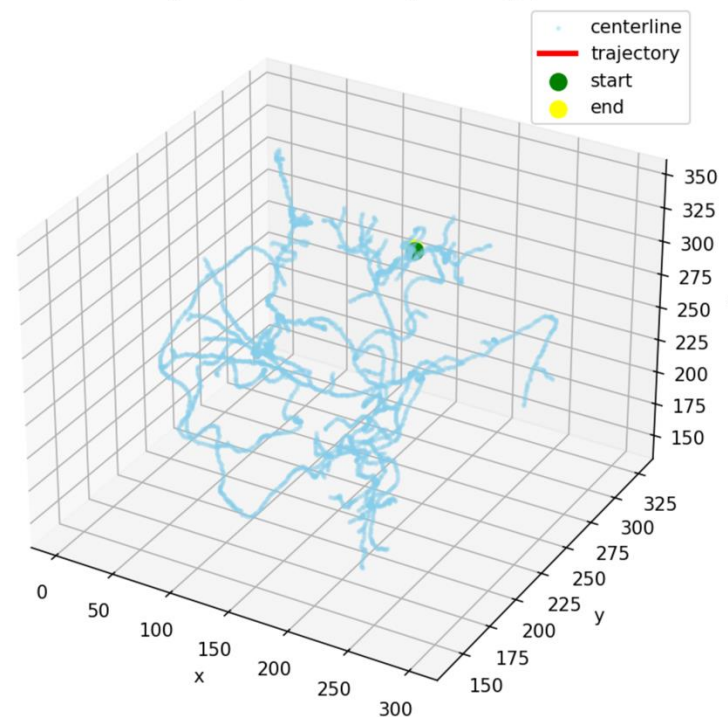
## 5. 결과 (test\_01)

- Trajectory 가 길게 생성 되지 않음.
  - 혈관을 exploration 하는게 아닌, 시작 주변 voxel 에서만 맵도는 현상
    - 특히 중앙의 굵은 대동맥 부분에서 두드러짐
  - 탐험행동에 더 높은 reward 고려

3D Trajectory vs Centerline | 01011ug\_529



3D Trajectory vs Centerline | 01010ug\_164



## 5. 실험 (test\_02)

- 모델구조 (v2)

- State (v2)

- 기존의 1채널 입력에서 2채널 입력으로 확장
      - ch0: 중심점 기준 patch (기존)
      - **ch1: visited\_mask (0 or 1)**
        - 탐험한 곳을 agent 가 알기 위한 주변 영역 brush 처리 (adaptive brush, 5x5x5 voxel) 후 state 에 재입력 (이 때, 실제 혈관영역만 칠해지도록 함 - [new\_pixel] 인자 생성)
        - 이동할 때마다 지속적으로 쌓아가는 구조

- Action (v2, 동일)

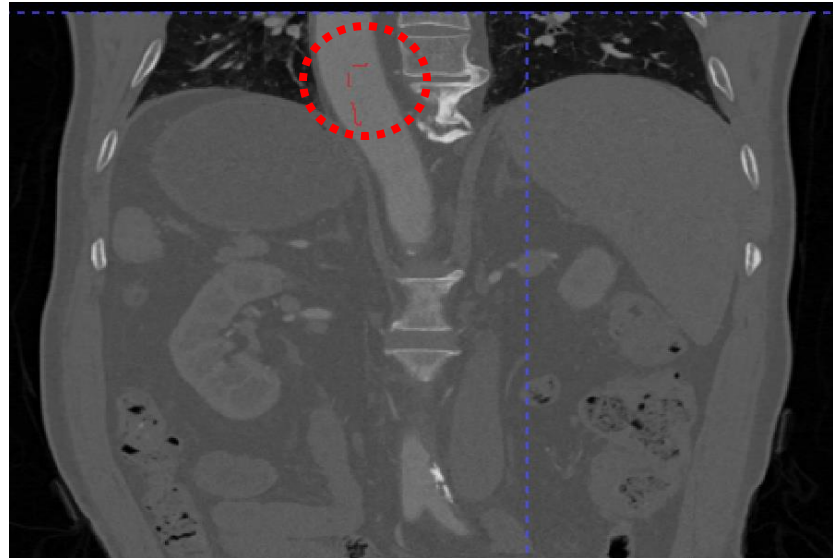
- 6 discrete action space

- Reward (v2)

- 혈관 밖 : -1.0 & episode 종료
  - **혈관 안**
    - If, [new\_pixel] > 0 : +1.0 + [new\_pixel \* 0.01]
    - Else, [new\_pixel] = 0 : -0.1

## 5. 결과 (test\_02)

- Trajectory 가 길게 생성 되지 않음.
  - 혈관을 exploration 하는게 아닌, 시작 주변 voxel 에서만 맵도는 현상
    - 특히 중앙의 굵은 대동맥 부분에서 탐험보상 때문에, 길게 왔다갔다 하는 현상
  - 혈관의 굵기에 따른 adaptive brush 의 부작용으로 생각됨
    - 굵기를 보정하는 parameter 필요



## 5. 실험 (test\_03)

- 데이터 전처리 [mask 데이터를 사용하여 centerline 추출]
  - Centerline 추출 library: skimage.morphology\_skeletonize
  - 실제 모델 사용 시에는 centerline 을 알수 없으므로, 보상에만 적용 되도록 함

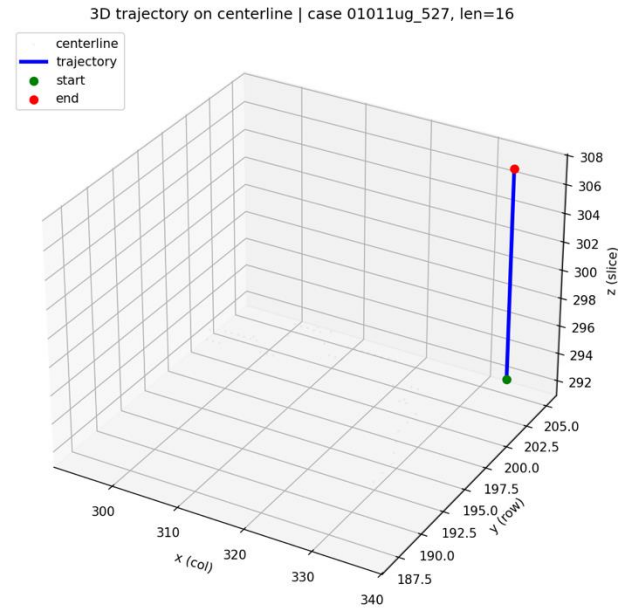
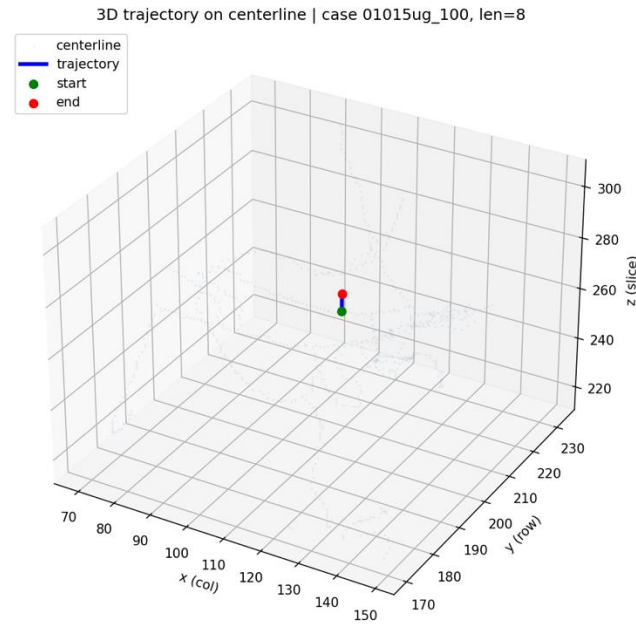
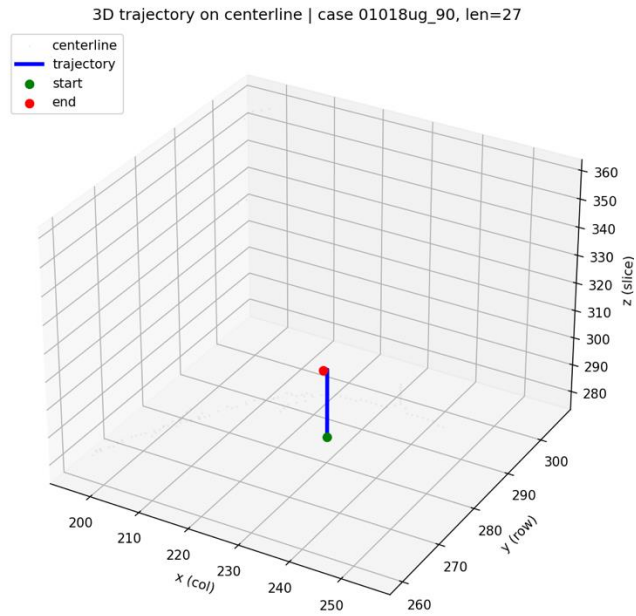
- 모델구조 (v3)

- State (v2, 동일)
  - ch0: 중심점 기준 patch (기존)
  - ch1: visited\_mask (0 or 1) (기존)
- Action (v2, 동일)
  - 6 discrete action space
- Reward (v3)
  - 혈관 밖 : -1.0 & episode 종료
  - 혈관 안
    - 브러쉬로 칠한 혈관 복셀 수에 따른 보상
      - $[\text{new\_pixel}] > 0 : +1.0 + [\text{new\_pixel} * 0.01]$
    - Centerline 방문 보상
      - 처음 방문한 centerline : +10
      - 이미 방문한 centerline : +0.1
      - Centerline 이 아닌곳 : -0.05



## 5. 결과 (test\_03)

- z 축으로만 움직임
  - Centerline 을 따라 가려는 행동은 전혀 하지 않음
- Centerline 의 복셀 방문은 space 에 비해 쉽지 않은 보상으로 생각됨



## 5. 실험 (test\_04)

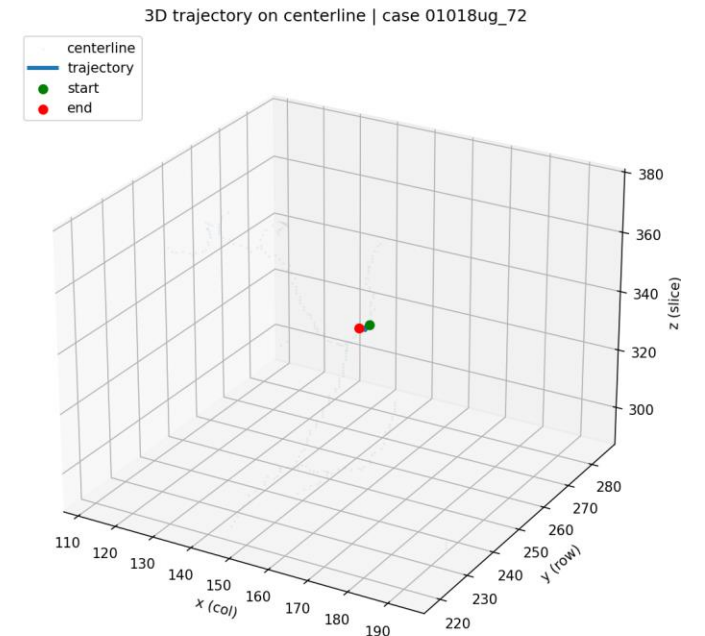
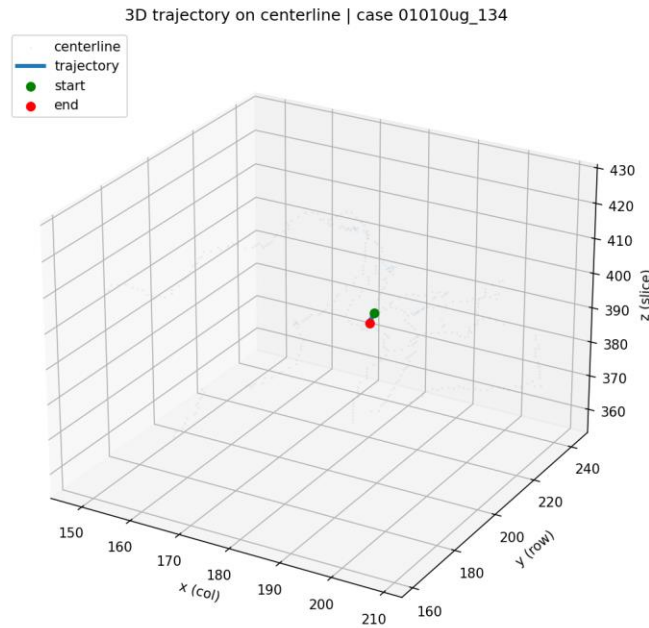
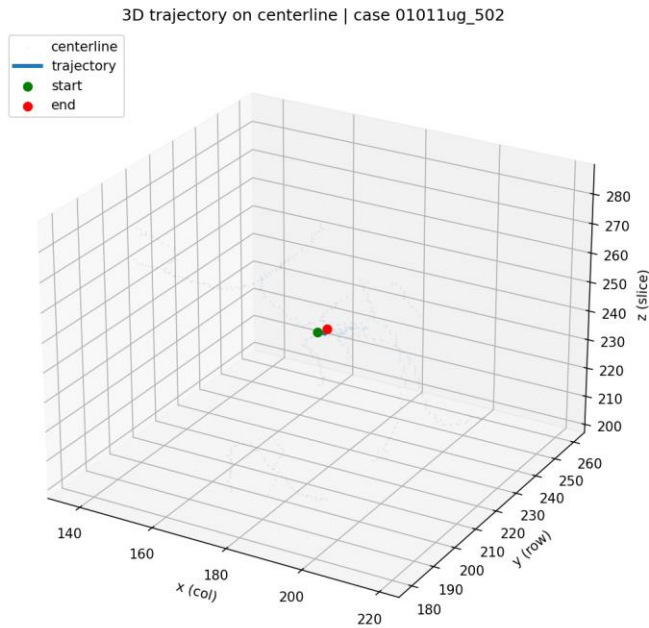
- Centerline 을 활용한 방향 계산 추가
  - 혈관의 centerline 방향으로 나아가는 것에 가중치를 준다
    - 기존 centerline 복셀 방문에 따른 가중치는 너무 harsh 함
    - Agent 가 나아가는 방향이 centerline 방향이면, 혈관을 따라 이동할 것이라 생각
  - Cos similarity 계산 추가
    - 이동 벡터 (t+1 에서의 좌표 - t 에서의 좌표)
    - Centerline 벡터 (t+1에서의 agent 좌표 주변 centerline의 최근접 복셀 5개의 SVD 를 통함)

- 모델구조 (v4)

- State (v2, 동일)
  - ch0: 중심점 기준 patch (기존)
  - ch1: visited\_mask (0 or 1) (기존)
- Action (v2, 동일)
  - 6 discrete action space
- Reward (v4)
  - 혈관 밖인 경우: -1.0 & episode 종료
  - 혈관 안인 경우
    - 브러쉬로 칠한 혈관 복셀 수에 따른 보상 (삭제)
    - Centerline 방향성 보상 (추가)
      - $\text{Reward} = + |\text{Cos similarity}|$
      - (방향성은 상관없으므로, 절대값 사용)

## 5. 결과 (test\_04)

- 혈관을 따라서 움직이지만 그 길이가 너무 짧음
  - Centerline 의 방향성 보상으로는 충분하지 않을 수 있음



## 5. 실험 (test\_05)

- Centerline 을 활용한 거리 계산 추가
  - 이동한 좌표 (t+1) 와 Centerline 과의 최소 거리에 따른 보상
- Reward scaling hyperparameter 추가

- **모델구조 (v5)**

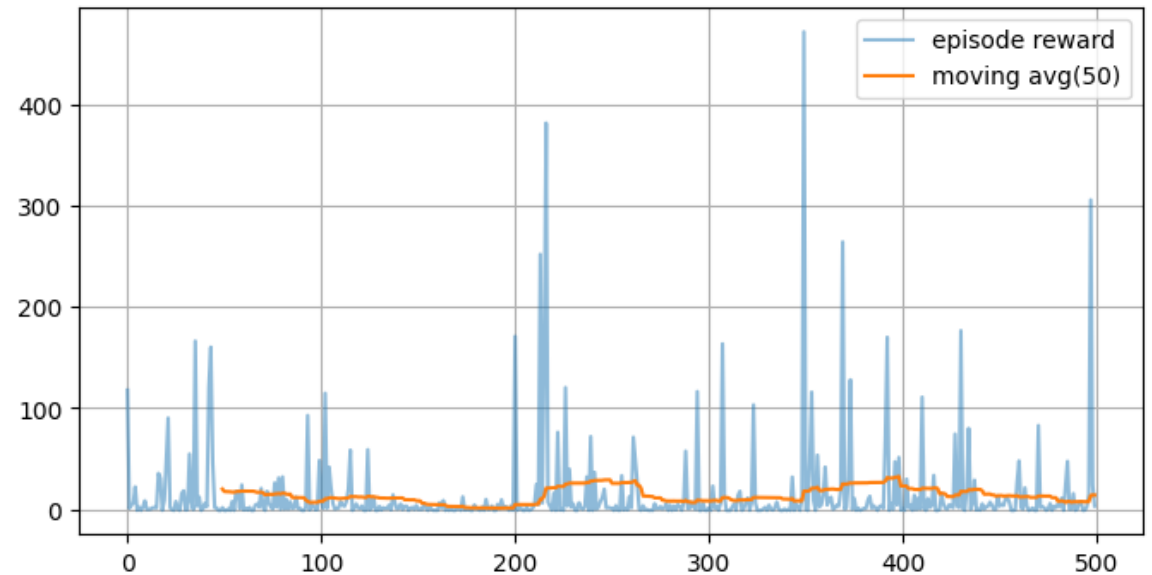
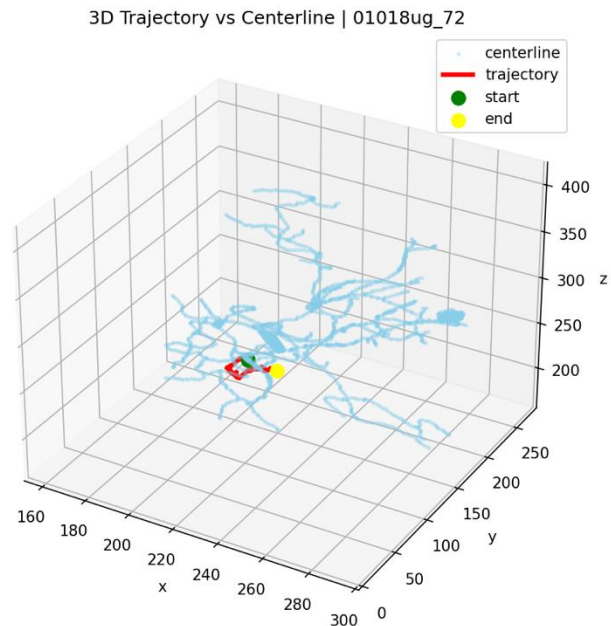
- State (v2, 동일)
  - ch0: 중심점 기준 patch (기존)
  - ch1: visited\_mask (0 or 1) (기존)
- Action (v2, 동일)
  - 6 discrete action space

- **Reward (v5)**

- 혈관 밖인 경우: -1.0 & episode 종료
- 혈관 안인 경우
  - Scailing parameter 적용 (추가)
    - $\text{Reward} = \text{Centerline 방향성 (기존)} * \text{방향성 hyperparameter} + \text{Centerline 과의 거리}(0 \sim 1 \text{ 정규화, 가까울 수록 } 1 \text{로}) * \text{거리 hyperparameter}$
  - Revisit penalty (추가)
    - Adaptive brush 로 칠해지는 혈관영역이 0 인 경우 : -0.5

## 5. 결과 (test\_05)

- 일부 validation 케이스에서 혈관을 따라서 이전 보다 길게 움직이지만 케이스 별 편차가 심함
  - 여전히 vessel 따라가는 탐험이 원할 하지 않다고 판단
- 학습 로그에서도 여전히 수렴하지 않는 듯 함



Training 시 episode 에 따른 reward 및 step

## 5. 실험 (test\_06)

- 모델구조 (v6)

- State (v2, 동일)
  - ch0: 중심점 기준 patch (기존)
  - ch1: visited\_mask (0 or 1) (기존)
- Action (v2, 동일)
  - 6 discrete action space

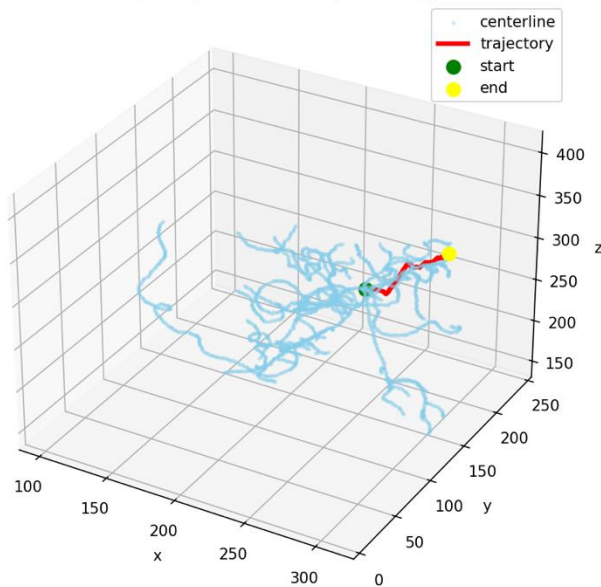
- Reward (v6)

- 혈관 밖인 경우: -1.0 & episode 종료
- 혈관 안인 경우
  - Centerline 방향성 보상(수정)
    - 기존 centerline 의 SVD 값이 특정 값 이하이면, 방향성 보상=0
    - 혈관의 분지 부분 등 방향 불명확 가능성 고려
  - Centerline 거리 보상 (기존)
  - 탐험 보상
    - 새 voxel 방문 에 따른 보상
      - Adaptive Brush = 0 : -0.5 (기존 revisit penalty)
      - Adaptive Brush < 5 : 0.1
      - Adaptive Brush < 20 : 0.3
      - Adaptive Brush > 20 : 0.5
  - 보상 총합
    - Reward =  
(Centerline 방향성 \* 방향성 hyperparameter) + (Centerline 과의 거리 \* 거리 hyperparameter) + 탐험보상

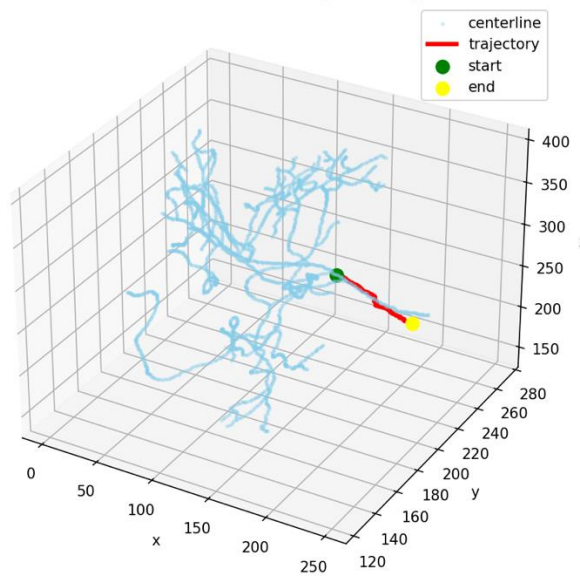
## 5. 결과 (test\_06)

- 일부 validation 케이스에서 혈관을 따라서 길게 움직이는 듯한 양상
- 일부 데이터에서 centerline 이 이상하게 출력되어 있음을 확인
  - Centerline 데이터에 대한 QC 누락 되었음

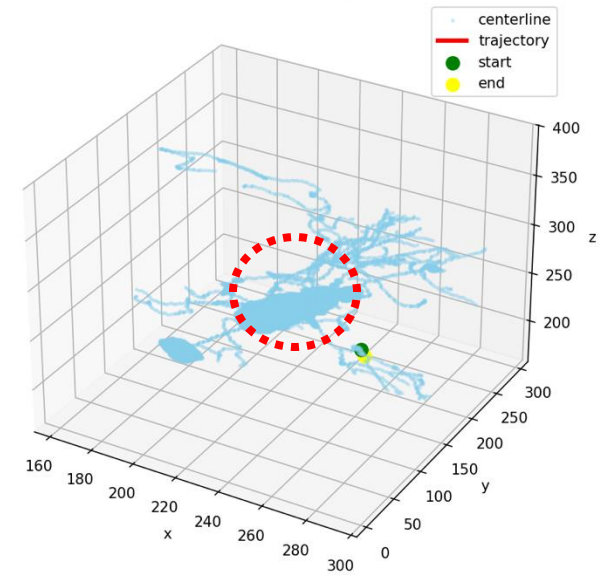
3D Trajectory vs Centerline | 01018ug\_86



3D Trajectory vs Centerline | 01010ug\_174



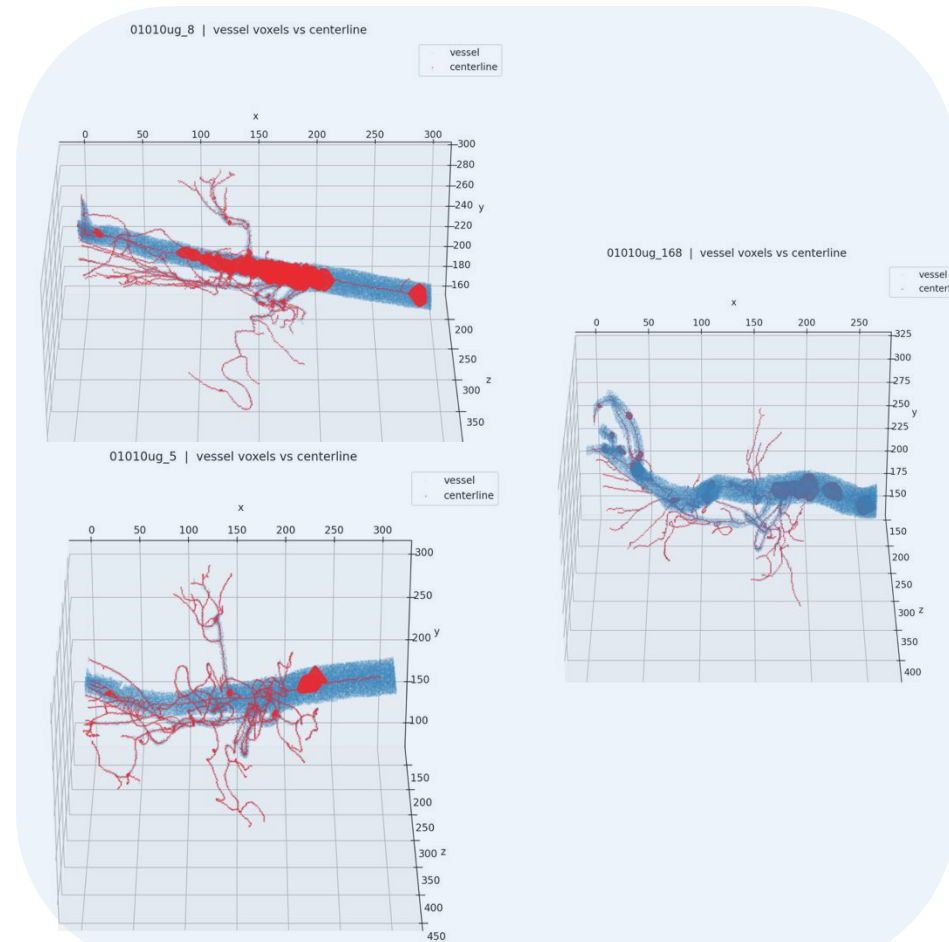
3D Trajectory vs Centerline | 01010ug\_8



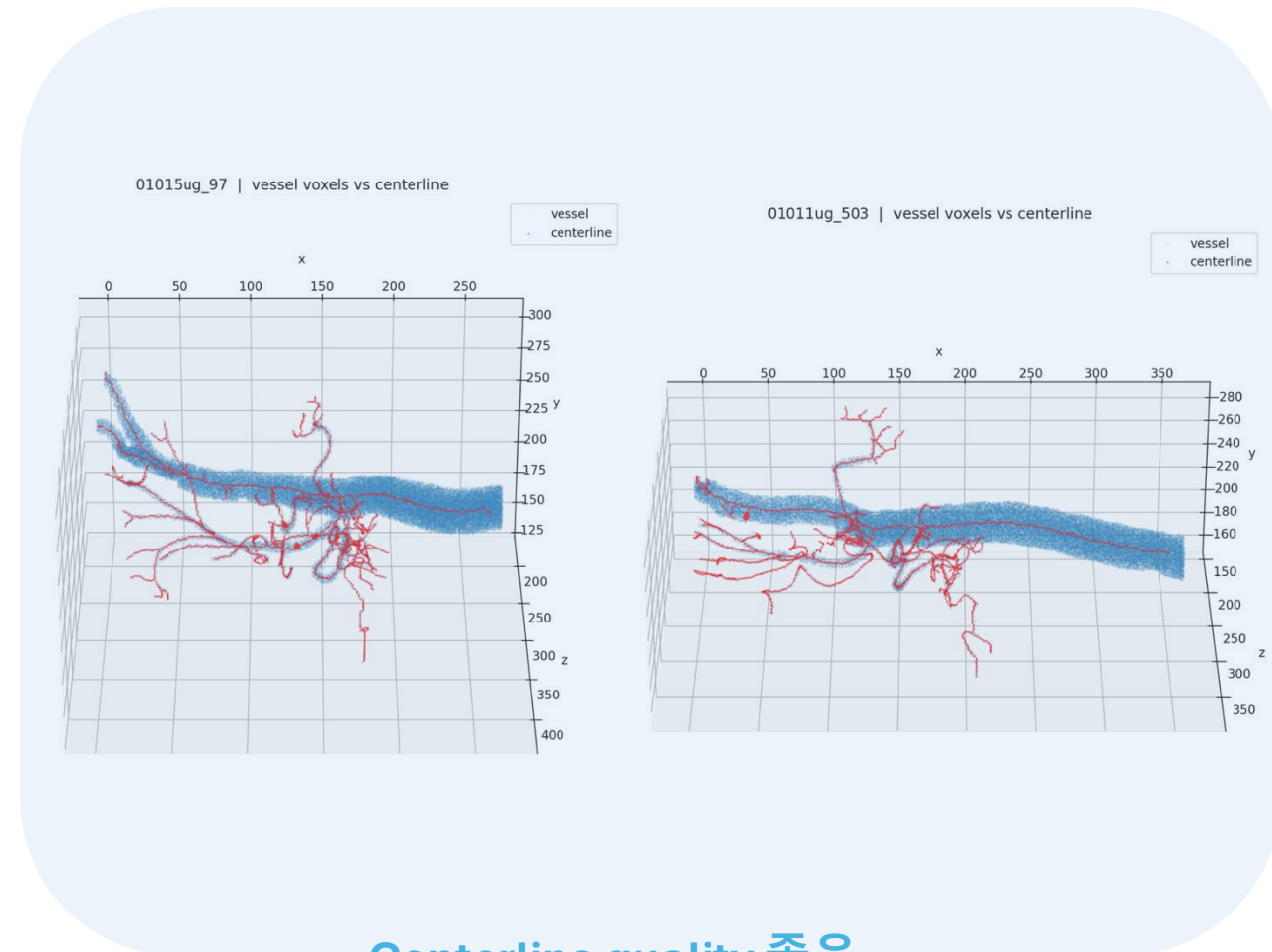


## 5. 실험 (test\_07) - preprocessing

- Centerline data QC 진행
  - 전체 99 케이스 중 centerline 이 깔끔하지 못한 케이스 20% 가량 있음



Centerline quality 좋지 않음

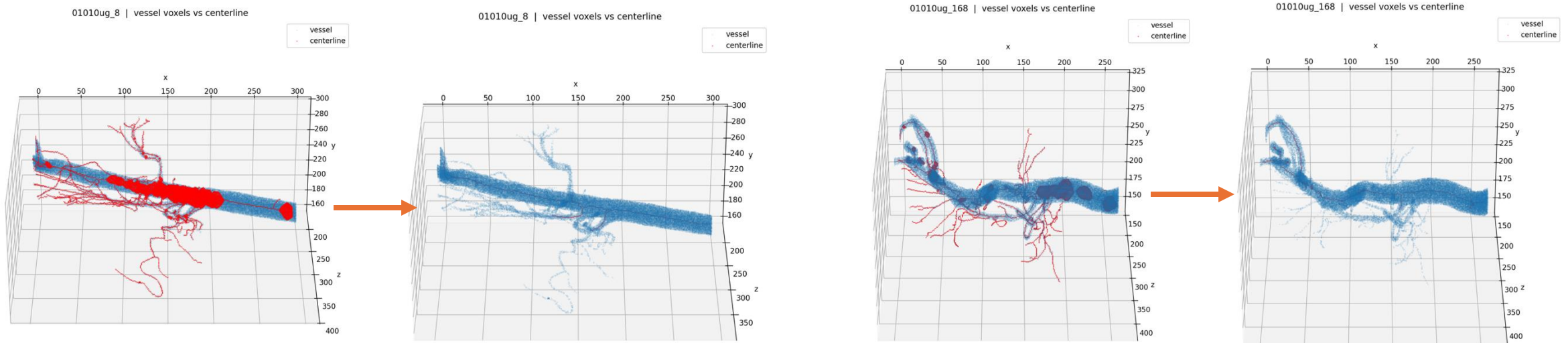


Centerline quality 좋음



## 5. 실험 (test\_07) - preprocessing

- Centerline 데이터 재생성
  - 기존 `skimage.morphology.skeletonize` -> simple ITK 의 `BinaryThinning` 기반으로 변경



- 기존의 얇은 혈관들의 centerline 은 다 사라지지만, 잘못된 centerline 으로 인한 reward 가 결과에 더 좋지 않다고 판단함
- Reward 의 재설계 필요

## 5. 실험 (test\_07)

- Centerline data 재생성
- Reward 재설계 – centerline 보다 탐험에 따른 보상 상승

- 모델구조 (v7)

- State (v2, 동일)
  - ch0: 중심점 기준 patch (기존)
  - ch1: visited\_mask (0 or 1) (기존)
- Action (v3)
  - 26 discrete action space
    - 현재위치를 중심으로 한 대각선 이동까지 고려

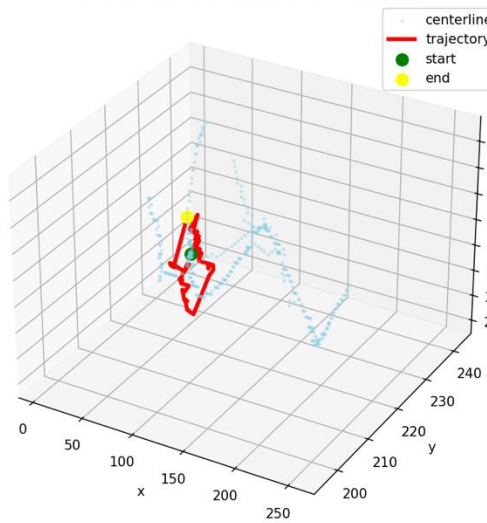
- Reward (v7)

- 혈관 밖인 경우: -1.0 & episode 종료
- 혈관 안인 경우
  - Centerline 방향성 보상(기존)
    - 기존 centerline 의 SVD 값이 특정 값 이하이면, 방향성 보상=0
  - Centerline 거리 보상 (삭제)
    - 얇은 혈관들의 centerline 은 없으므로 삭제
  - 탐험 보상 (scale 대폭 상승)
    - 새 voxel 방문 에 따른 보상
      - Adaptive Brush = 0 : -0.5 (기존 revisit penalty)
      - Adaptive Brush < 5 : 0.5
      - Adaptive Brush < 20 : 0.8
      - Adaptive Brush > 20 : 1.0
- 보상 총합
  - $\text{Reward} = (\text{Centerline 방향성} * \text{방향성 hyperparameter}) + \text{탐험보상}$

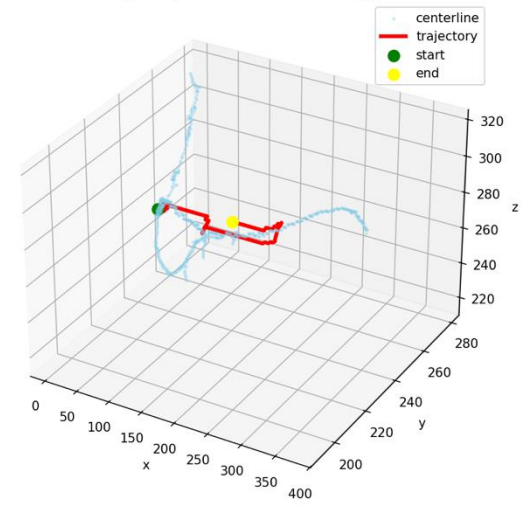
# 5. 결과01 (test\_07)

- Centerline 과 trajectory 시각화 (3D plot)
  - validation 케이스들 에서 혈관을 따라서 이전보다는 훨씬 길게 움직이는 듯한 양상
  - Centerline 의 재생성으로 맞게 갔는지 다른 시각화 방식 필요

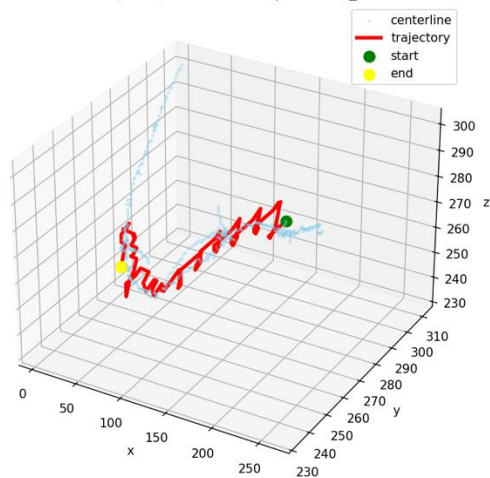
3D Trajectory vs Centerline | 01010ug\_174



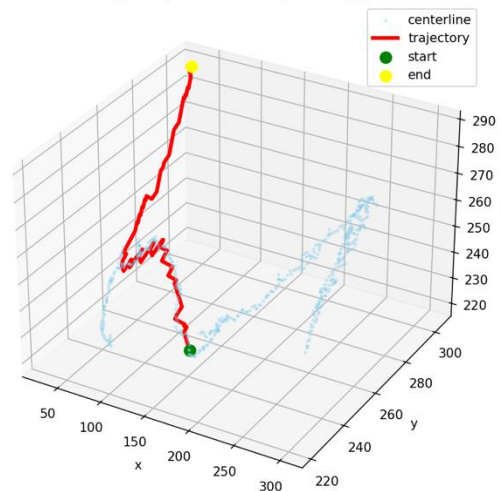
3D Trajectory vs Centerline | 01011ug\_502



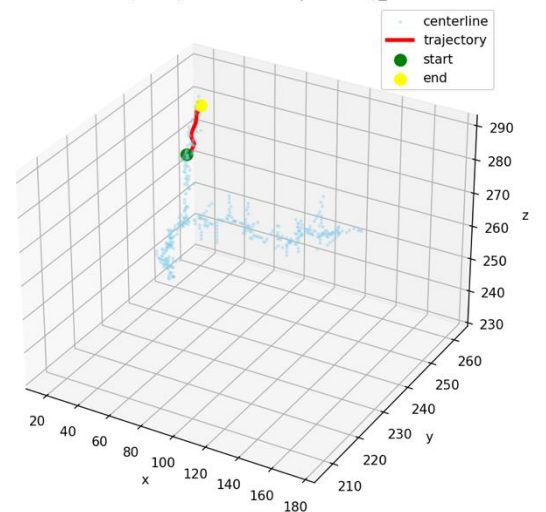
3D Trajectory vs Centerline | 01018ug\_85



3D Trajectory vs Centerline | 01010ug\_164

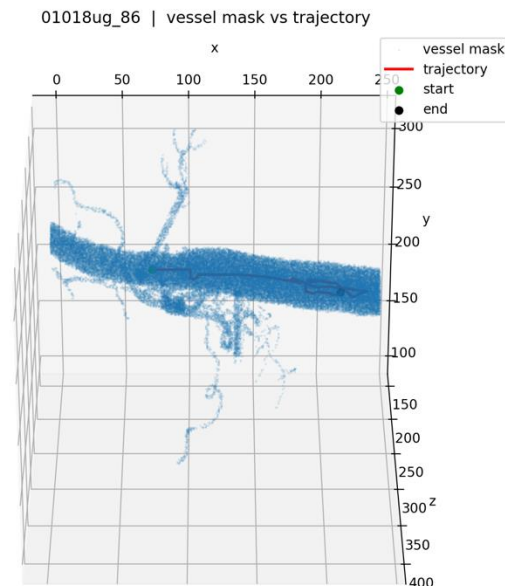
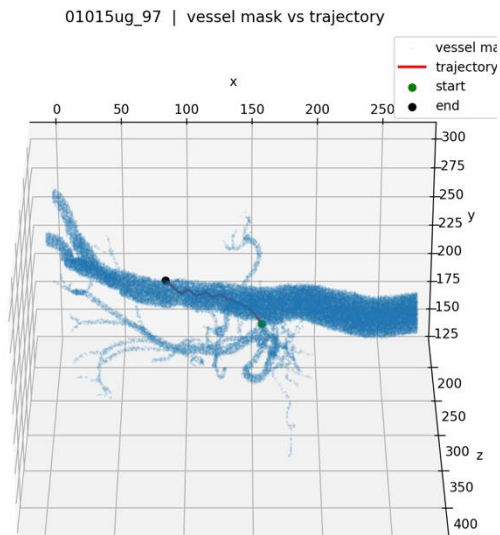
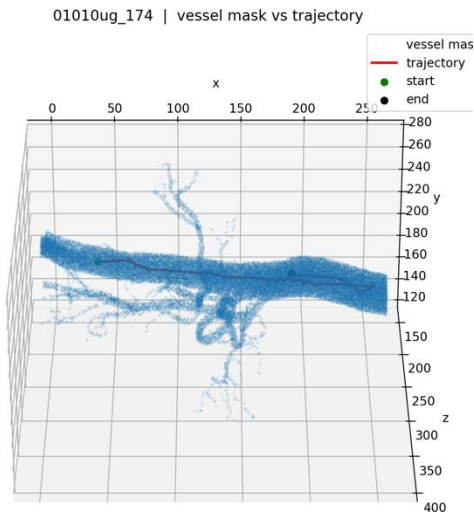
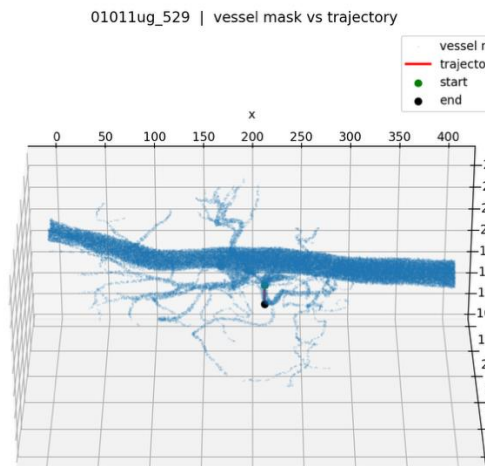


3D Trajectory vs Centerline | 01015ug\_97



## 5. 결과02 (test\_07)

- 혈관 mask 와 trajectory 시각화 (3D plot)
  - 굵은 혈관인 대동맥을 따라 길게 찾아감
    - 이전 보다 이동 거리가 길어 졌으나, 여전히 한계
    - 시작점의 위치가 혈관 voxel 중 random 이라 대부분은 대동맥 위에서 학습하였을 것임
  - 얇은 혈관에서 운 좋게 시작한 agent 에서는 얼마 못가 out\_of\_vessel 로 episode 종료 됨
    - 3D CNN 의 end to end 학습 방식이 back propagation 이 원활하지 않을 수 있음
      - 기존 의료영상 segmentation 표준 알고리즘 을 encoder 로 사용하는 것이 나을 수 있음



## 6. 결과 종합 및 개선 고려사항

- 혈관의 두께에 따른 reward 고려 실패
  - 대동맥에서만 여전히 길게 움직이려 함
  - Centerline 에 대한 정확도 미리 확인 하지 못함
- Agent 의 시작점 설계에 따른 학습 불균형
  - 다양한 시작점에서 학습을 고려 하거나, episode 를 충분히 길게 설계
- 다중 agent 환경고려 필요
  - 혈관의 분지점에서 agent 의 추가가 필요
- 실제 적용 고려 시
  - 현재의 centerline tracking 후에 segmentation 이 될수 있는 모델 병합 필요