

Efficient Pix2Vox++ for 3D Cardiac Reconstruction from 2D echo views

<https://arxiv.org/pdf/2207.13424>

0. Introduction

- 표준 2D 심장 초음파(echo)는 3D 해부학적 구조 정보를 직접 제공하지 못함
- 3D 심장 구조는 심장 질환 진단, 용적 측정, 기능 평가에 필수적
- 기존 2D→3D 재구성 모델들은 메모리 사용량이 크고 의료 영상 적용이 어려움
- 기존 연구들은 untracked 2D echo view로부터 전체 3D 심장을 복원한 사례가 거의 없음
- 본 연구는 표준 2D echo 뷰만으로 3D 심장 구조를 재구성하는 효율적 모델(E-Pix2Vox) 을 제안함

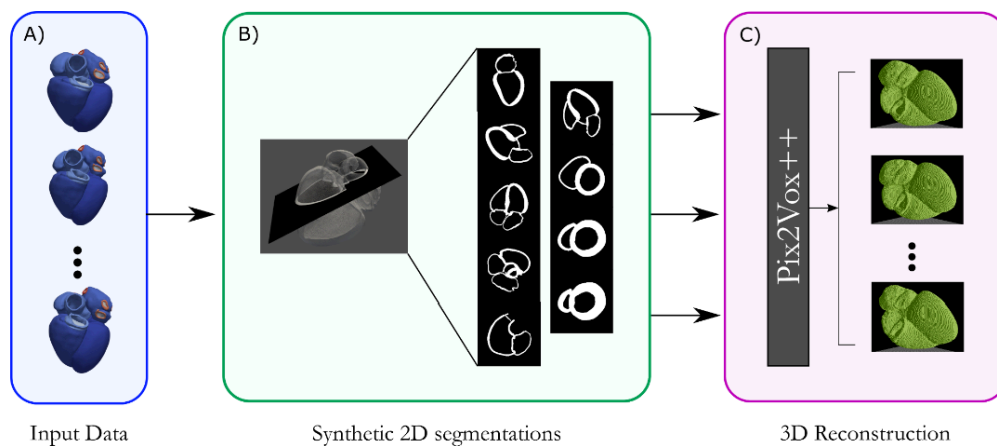
1. Overview

- 여러 개의 2D echo 뷰를 입력으로 받아 하나의 3D 심장 볼륨을 복원하는 구조
- 기존 Pix2Vox++ 구조를 단순화한 E-Pix2Vox(E-PiVox) 제안
- 다중 뷰 latent를 통합한 후 단일 decoder로 3D volume 복원
- synthetic cardiac mesh 기반 데이터 생성 후 학습
- segmentation mask 입력과 ultrasound-like 이미지 입력 모두 실험
- 심장 초음파 기반 3D reconstruction의 임상적 활용 가능성 제시

2. Challenges

- 심장은 벽이 얇고 곡률이 큰 복잡한 3D 구조
- 2D 뷰만으로는 깊이 정보 소실 발생
- 초음파 영상 특유의 노이즈, 음영, 아티팩트로 인한 해석 어려움
- real-world 2D echo + 3D GT 데이터 부족
- synthetic 데이터 기반 학습 → domain gap 문제 존재
- 기존 Pix2Vox 계열 모델은 메모리 사용량이 매우 큼

3. Method



- 기존 Pix2Vox++의 다중 decoder 병렬 구조 제거
- 모든 뷰를 latent space로 압축 후 **단일 decoder로 3D 복원**
- 각 2D segmentation mask를 encoder에 통과시켜 feature 추출
- latent feature들을 3D convolution으로 병합
- synthetic 3D 심장 mesh → standard view slicing → 2D mask 생성
- mask → ultrasound-like 이미지 변환에 CycleGAN 사용
- Adam optimizer 사용, 200 epoch 학습

4. Experiments

- 데이터셋

- 1000개의 synthetic 3D cardiac mesh
- train 70% / val 15% / test 15%
- 입력 유형
 - segmentation mask 기반 2D view
 - ultrasound-like 이미지 기반 2D view
- 비교 모델
 - Pix2Vox++ Fast / Accurate
 - E-Pix2Vox Fast / Accurate
- 평가 지표
 - IoU (Intersection over Union)
- 실험 설정
 - view 개수 1, 2, 4, 8, 9에 따른 성능 비교

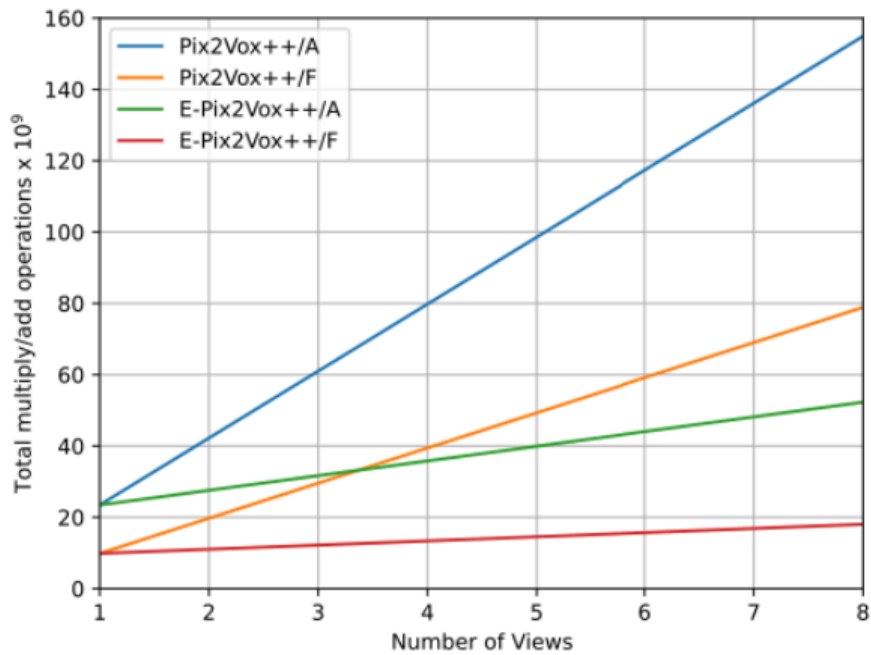
5. Results

Table 2. IoU Comparison of multi-view 2D to 3D object reconstruction on the ShapeNet data set for a 32^3 resolution.

Method	1 View	2 Views	4 Views	8 Views
PiVox/Fast	0.645	0.669	0.682	0.690
E-PiVox/Fast	0.641	0.663	0.675	0.682
Percentage diff.(%)	-0.620	-0.897	-1.026	-1.16
PiVox/Accurate	0.670	0.695	0.708	0.715
E-PiVox/Accurate	0.6687	0.6836	0.6949	0.701
Percentage diff.(%)	-0.19	-1.64	-1.85	-1.96

Table 3. IoU Comparison of multi-view 2D to 3D object reconstruction on the cardiac data set for a 64^3 resolution. A2C and A4C refer to atrial 2 and 4 chamber image views respectively, and 9 views denotes all 9 of the previously described standard ultrasound views.

method	Binary				CycleGAN inferenced		
	9 views	A2C	A4C	A2C & A4C	A2C	A4C	A2C & A4C
PiVox/Fast	0.859	0.828	0.841	0.818	0.644	0.652	0.7
E-PiVox/Fast	0.86	0.837	0.85	0.822	0.651	0.66	0.712
PiVox/Accurate	0.895	0.85	0.867	0.873	0.673	0.732	0.741
E-PiVox/Accurate	0.903	0.868	0.878	0.881	0.674	0.734	0.74



- ShapeNet 객체 재구성 실험
 - E-Pix2Vox는 Pix2Vox 대비 IoU 성능 저하가 매우 작음 (약 1~2%)
- 심장 재구성 (segmentation mask 입력)
 - 9 views: IoU \approx 0.903
 - 2 views: IoU \approx 0.881
- 심장 재구성 (ultrasound-like 입력)
 - 2 views: IoU \approx 0.741
- 입력 뷰 수 증가 \rightarrow 재구성 품질 상승
- 얇은 심장 벽, 밸브 영역에서 일부 복원 오류 발생

6. Insight

- 적은 수의 2D echo 뷰만으로도 전체 3D 심장 구조 복원이 가능함을 입증
- E-Pix2Vox 구조는 **의료 영상용 3D reconstruction에서 효율성 측면의 실질적 이점**을 제공
- synthetic 데이터 + CycleGAN 전략은 의료 영상 데이터 부족 문제에 대한 현실적 접근

- real echo 데이터 일반화 성능은 아직 불확실
- 병변이 있는 심장, 비정상 구조에 대한 검증 필요
- 향후 real-world 3D GT 포함 학습, temporal(4D) heart reconstruction 연구로 확장 가능