Anatomy Completor: A Multi-class Completion Framework for 3D Anatomy Reconstruction

Jianning Li[1], Antonio Pepe[2], Gijs Luijten[1,2], Christina Gsaxner[2], Jens Kleesiek[1], Jan Egger[1]

[1]Institute for Artificial Intelligence in Medicine (IKIM), University Hospital Essen (AöR), Essen, Germany
[2]Institute of Computer Graphics and Vision (ICG), Graz University of Technology, Graz, Austria
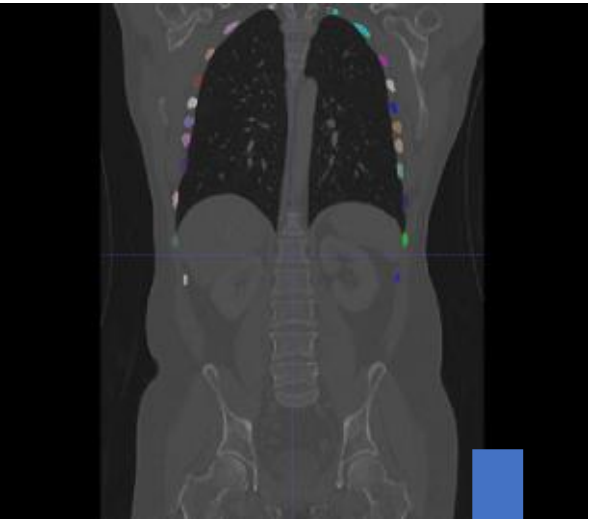
contact: jianning.li@uk-essen.de

Problem Statement

For an unannotated CT dataset, given partial manual annotations, how to automatically generate whole-body (pseudo) annotations?
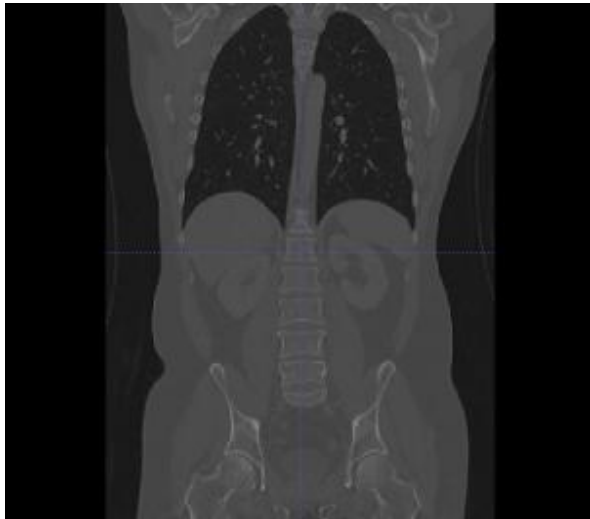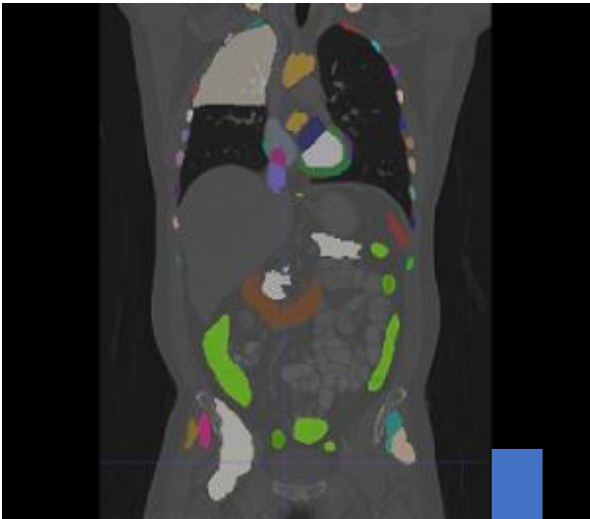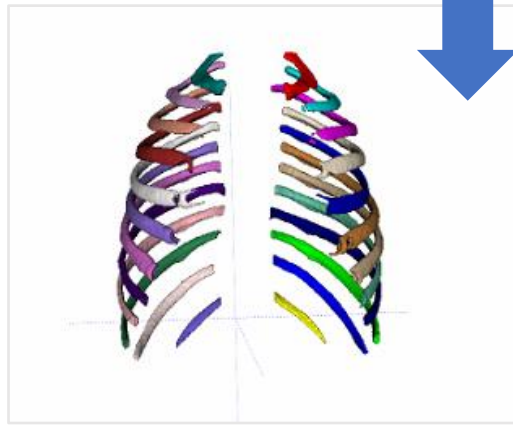
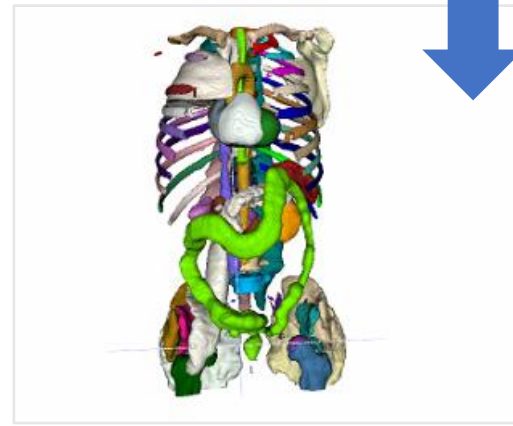CT scan   sparse/partial annotations   whole-body (pseudo) labels

3D shape completion (image from [1])
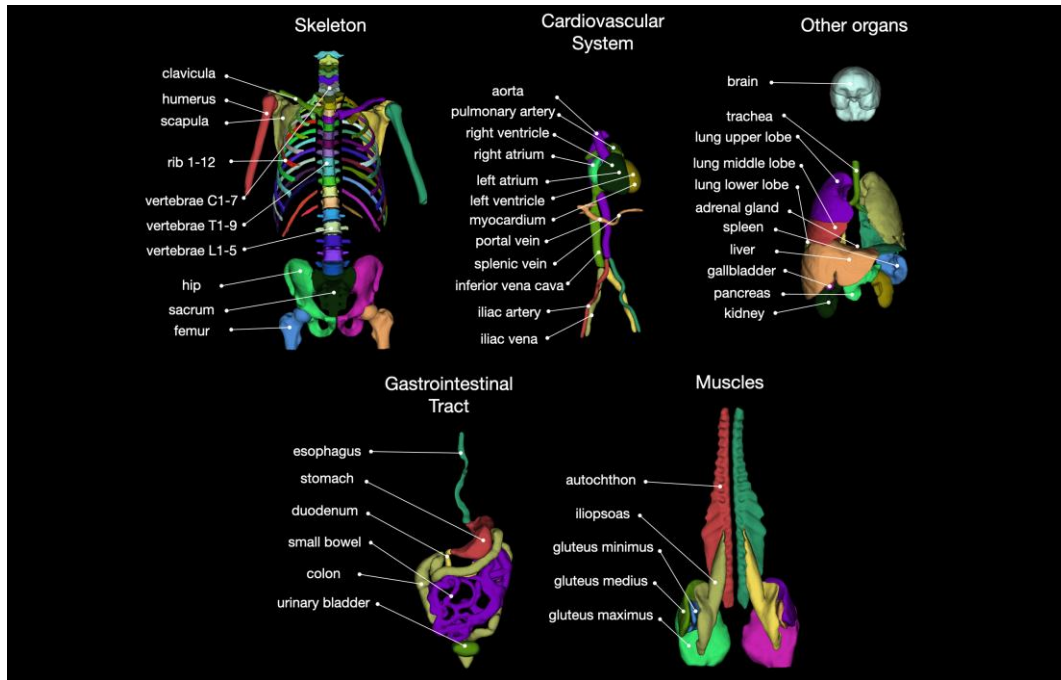
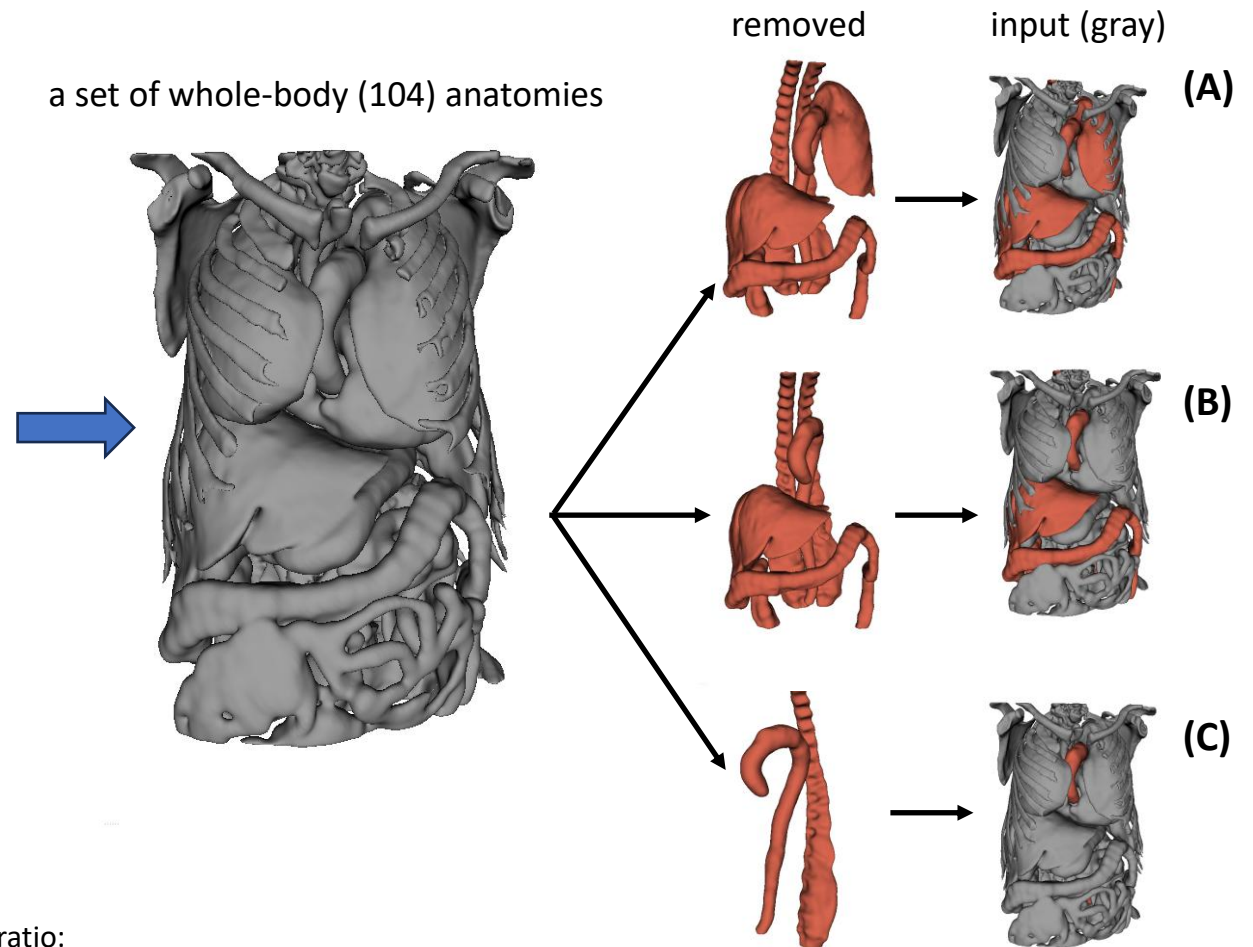input: partial labels   target: whole-body labels

[1] Chu, L., et al. *Unsupervised shape completion via deep prior in the neural tangent kernel perspective*. ACM Transactions on Graphics (TOG), 2021

Dataset Creation: Single Class Dataset

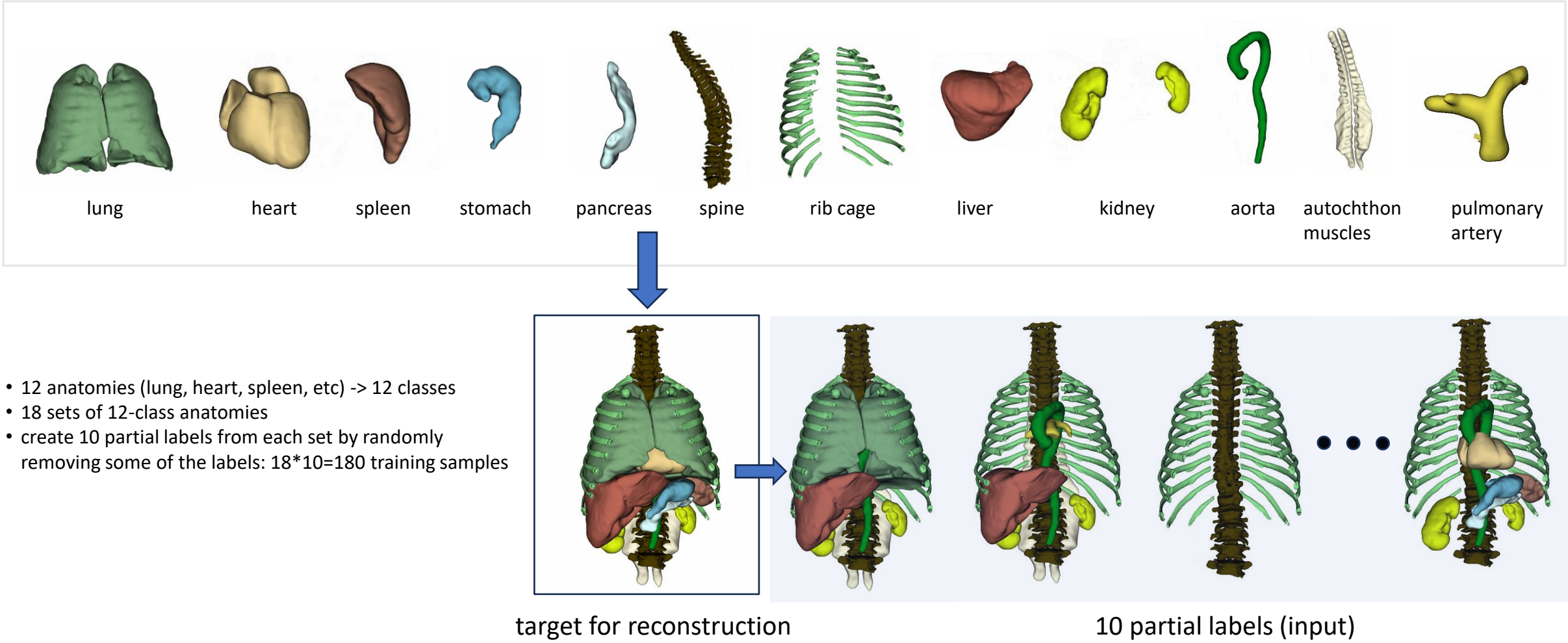Totalsegmentor whole-body segmentations [1]

a set of whole-body (104) anatomies

removed   input (gray)

(A)

(B)

(C)

- 737 sets of whole-body anatomies (451 training, 286 evaluation)
- Each set contains 104 anatomies
- To create partial labels, remove some of the anatomies based on their volume ratio:
  - anatomies with a volume ration of **10%** and above are removed **(A)**
  - anatomies with a volume ration of **20%** and above are removed **(B)**
  - anatomies with a volume ration of **40%** and above are removed **(C)**

Single-class dataset: anatomies are not distinguished (different anatomies have the same label "1". The background is "0")

[1] Wasserthal, J., et al., **TotalSegmentator: robust segmentation of 104 anatomical structures in CT images**. arXiv preprint arXiv:2208.05868 (2022)

Dataset Creation: Multi-class Dataset

lung   heart   spleen   stomach   pancreas   spine   rib cage   liver   kidney   aorta   autochthon muscles   pulmonary artery

- 12 anatomies (lung, heart, spleen, etc) -> 12 classes
- 18 sets of 12-class anatomies
- create 10 partial labels from each set by randomly removing some of the labels: 18*10=180 training samples

target for reconstruction

10 partial labels (input)

Method: 3D Denoising Auto-encoder (DAE)

- Learning a **many-to-one** mapping

$$\mathcal{F} : \{x_n^m\}_{m=1}^{M} \to y_n, \; n = 1, 2, ..., N$$

- Learning a **one-to-one (residual)** mapping

$$\mathcal{F}_{res} : \{x_n^m\}_{m=1}^{M} \to \{y_n - x_n^m\}_{m=1}^{M}, \; n = 1, 2, ..., N$$

- Loss function: two variants of **Dice** loss

many-to-one mapping   residual mapping

$$\mathcal{L}_{\mathcal{F}} = \sum_{m=1}^{M} \sum_{n=1}^{N} \mathcal{L}_{dice}(y_n, \tilde{y}_n^m) \quad \mathcal{L}_{\mathcal{F}_{res}} = \sum_{m=1}^{M} \sum_{n=1}^{N} \mathcal{L}_{dice}(y_n, \tilde{x}_n^m + x_n^m)$$
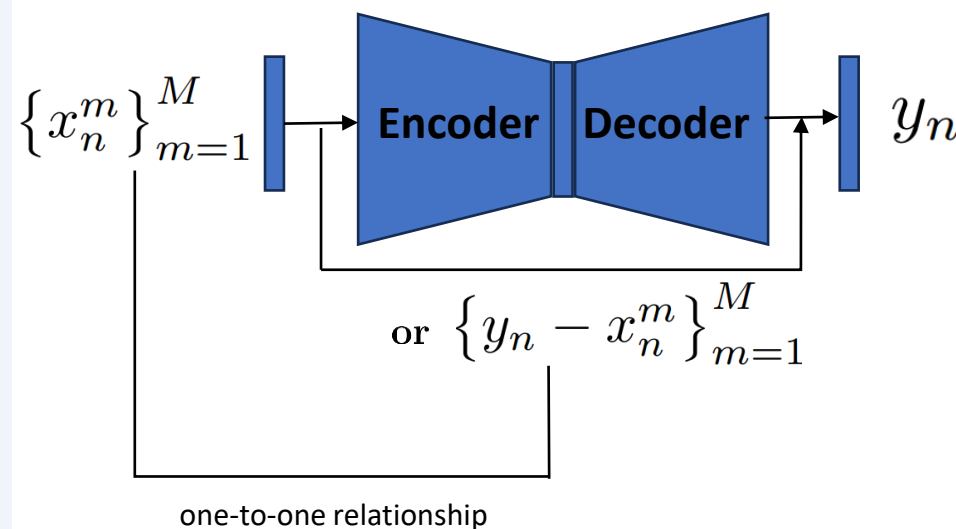
x_n^m the *m*th partial label from set *n*

y_n full label from set *n*

$\{y_n - x_n^m\}_{m=1}^{M}$ the missing labels of set *n*

$\mathcal{L}_{dice}$ Dice loss

- a one-to-one mapping is easier to learn than a many-to-one mapping
- DAE can detect what are missing in the input (which are random), and reconstruct them. DAE is expected to learn every possible combinations.
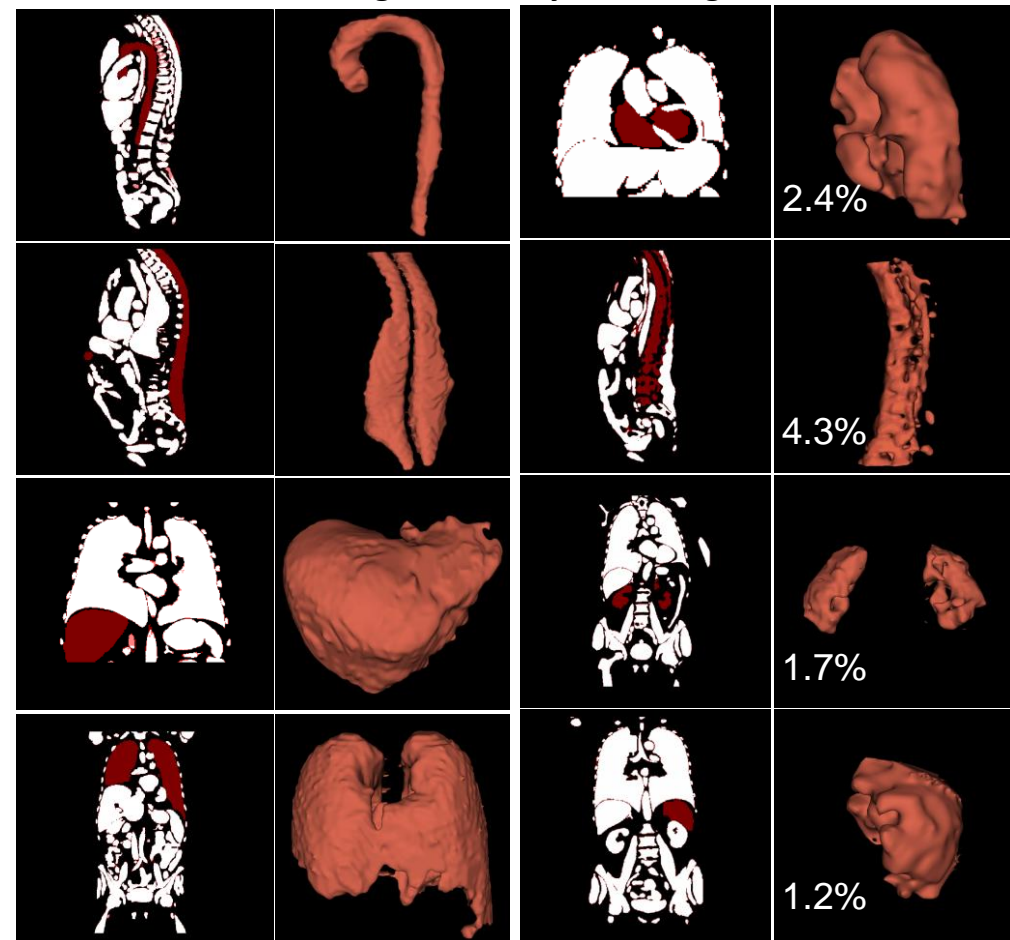- increasing *M* (the number of partial labels) increases the DAE's ability to do so

$$\{x_n^m\}_{m=1}^{M} \to \boxed{\text{Encoder} \;|\; \text{Decoder}} \to y_n$$

or $\{y_n - x_n^m\}_{m=1}^{M}$

one-to-one relationship

Results: Single-class Reconstruction (104 anatomies, 1 class)

multiple anatomies are missing
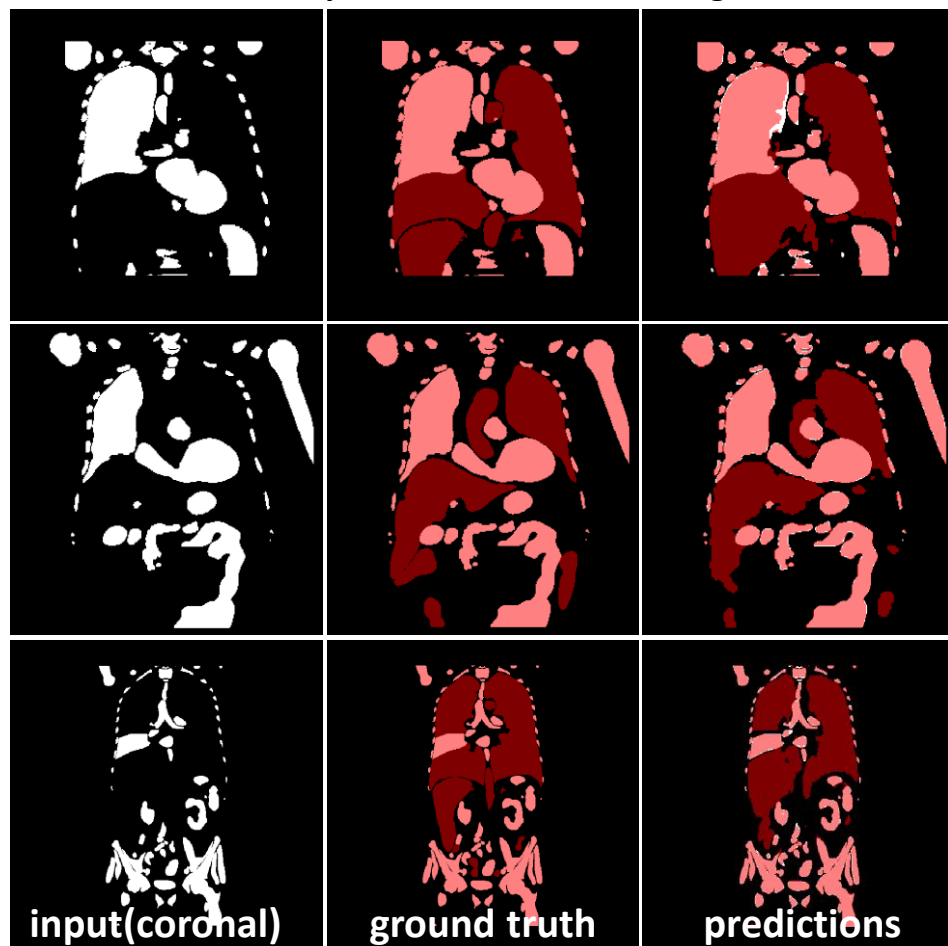
a single anatomy is missing

input(coronal) ground truth predictions

2.4%
4.3%
1.7%
1.2%
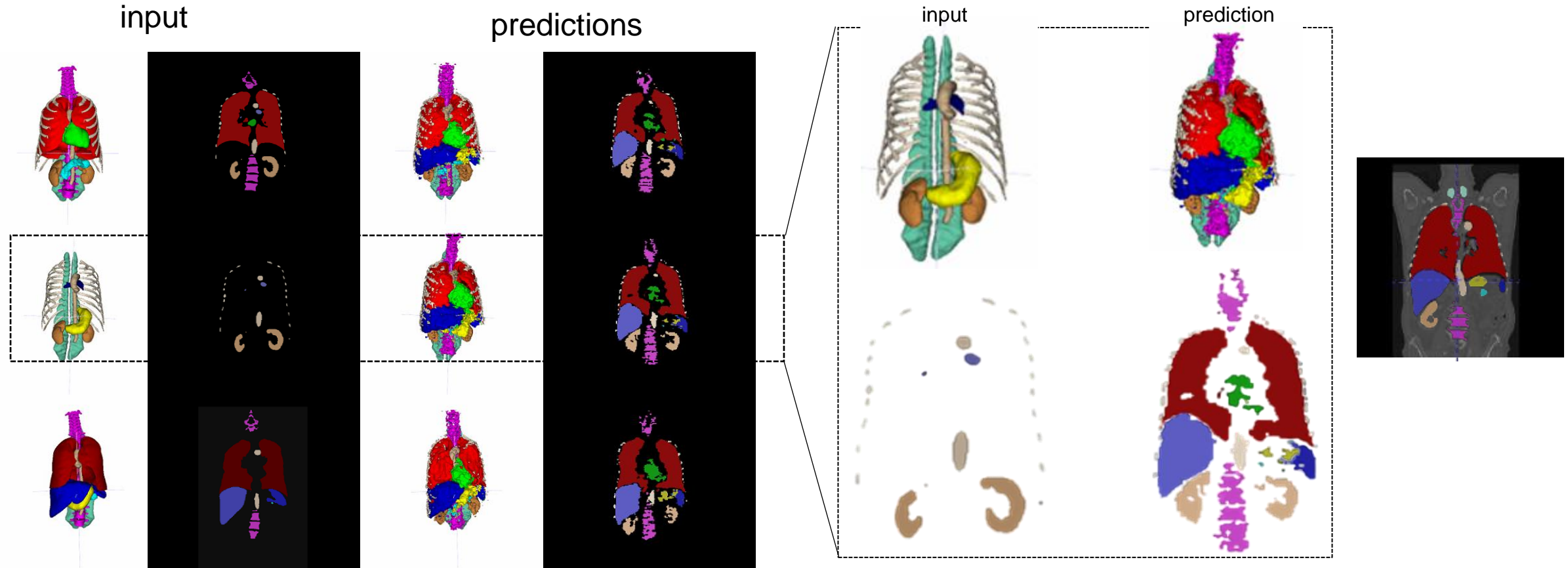
- on three test sets D_{test1} (10%), D_{test2} (20%), D_{test3} (40%), the DSC (standard deviation) are **0.865 (0.074), 0.904 (0.039), 0.931 (0.030)**
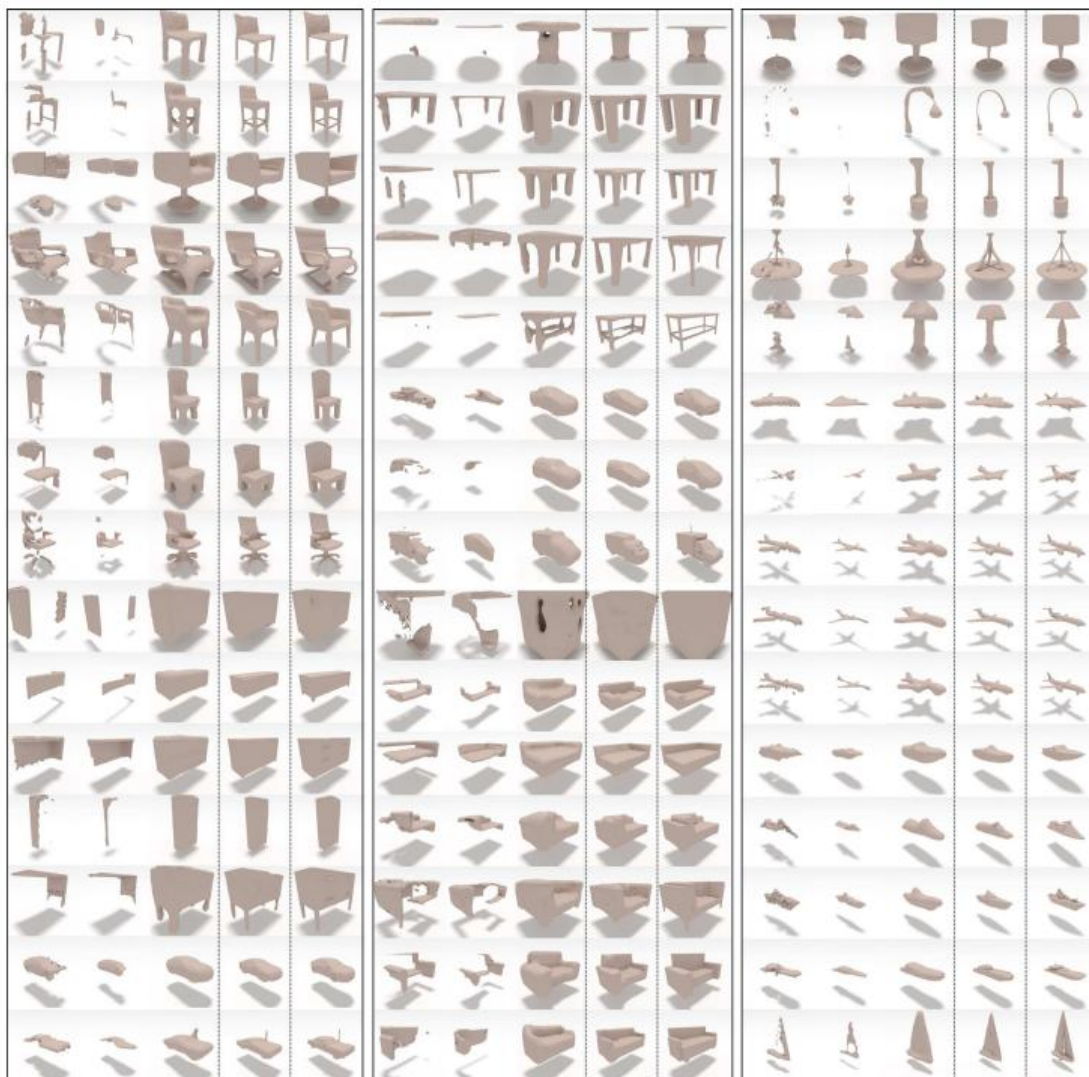- better performance in reconstructing larger missing anatomies (D_{test3} 40%)

Results: Multi-class Reconstruction (12 anatomies, 12 classes)

input   predictions   input   prediction

- Generate whole-body pseudo labels for an unannotated dataset (to train a whole-body segmentation network)
- Annotate the rib cage (and a few anchoring organs), and generate pseudo labels for the remaining anatomies
- How to choose the proper anchors, so that the positions of the reconstructed anatomies match the corresponding CT scan the best?
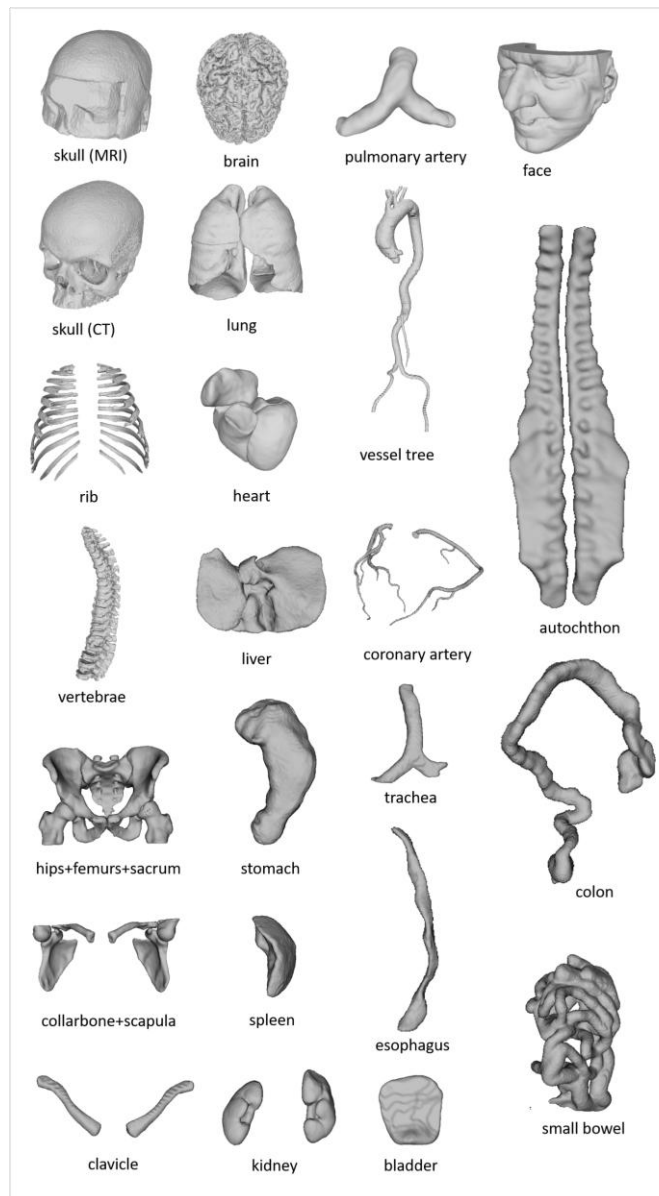
MedShapeNet

MedShapeNet: (1) 3D models of normal and pathological human anatomies extracted from the imaging data of real patients. (2) a medical version of *ShapeNet.*
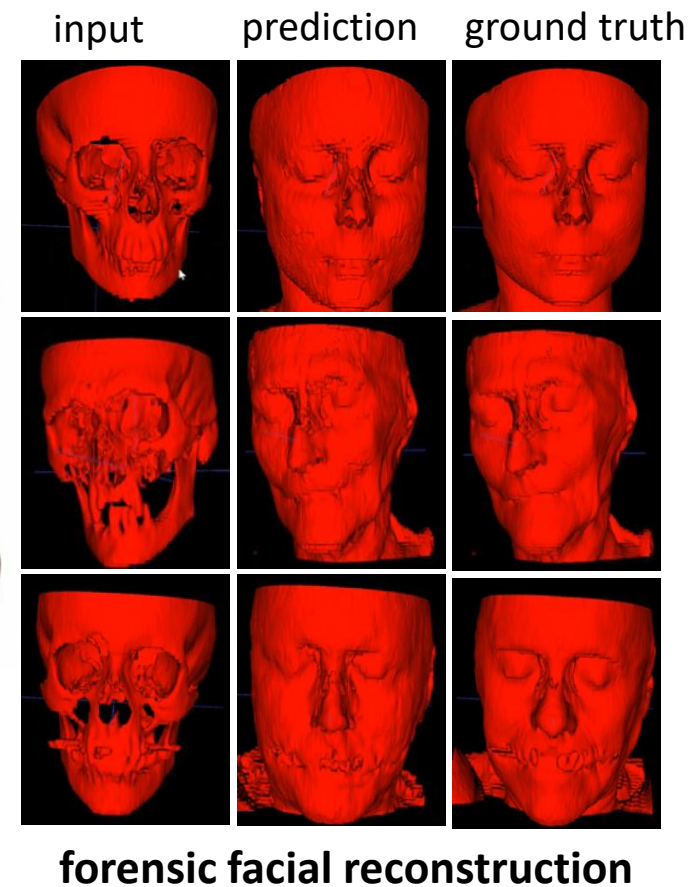
ShapeNet: 3D **CAD models** of **real-world objects**: chair, desk, car, airplane…

benckmark dataset: 3D shape completion, retrieval/classification, reconstruction…

https://shapenet.org/

skull (MRI)   brain   pulmonary artery   face

skull (CT)   lung

rib   heart   vessel tree

vertebrae   liver   coronary artery   autochthon

hips+femurs+sacrum   stomach   trachea   colon

collarbone+scapula   spleen   esophagus   small bowel

clavicle   kidney   bladder

normal

tumorous (paired) kidneys

tumorous brains

diseased coronary arteries

tumorous livers

defected skull   head & neck tumors

pathological

https://medshapenet.ikim.nrw/

MedShapeNet: other shape reconstruction applications

input prediction groundtruth

input prediction ground truth

facial bone reconstruction

cranium bone reconstruction

forensic facial reconstruction
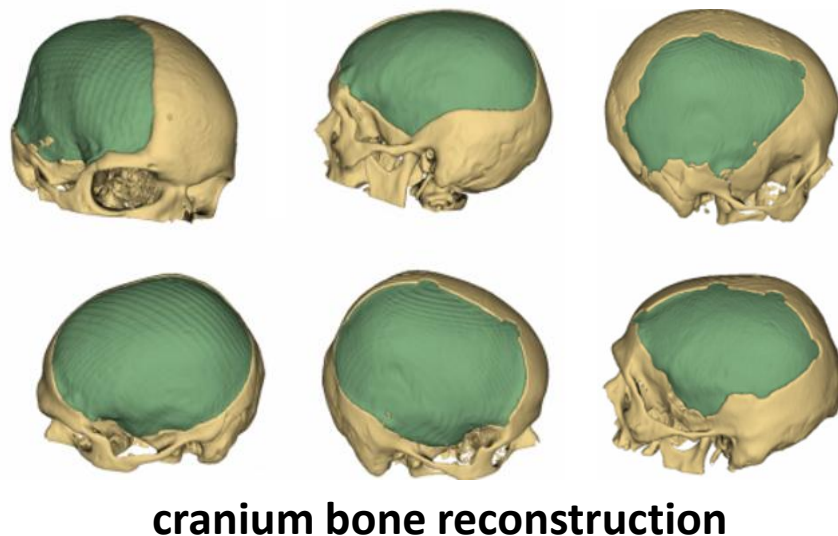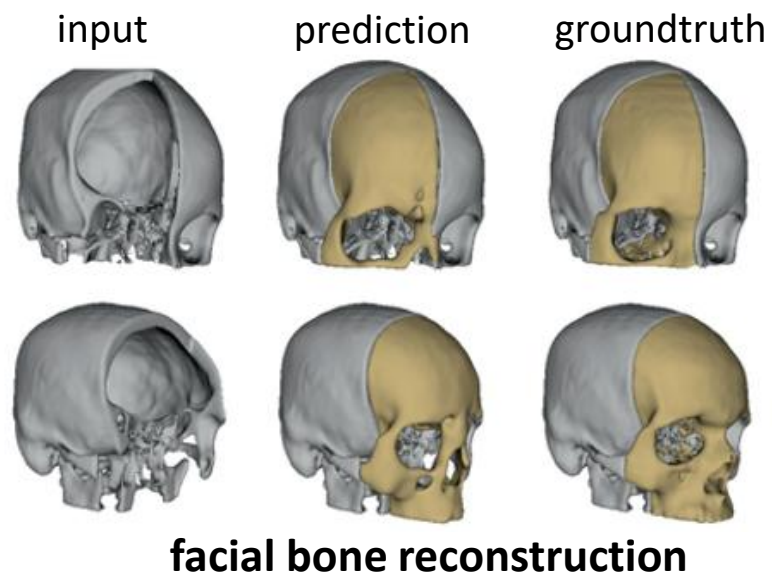
benckmark datasets and codes:
https://github.com/Jianningli/medshapenet-feedback/

Li, J., et. al., 2023. *MedShapeNet - A Large-Scale Dataset of 3D Medical Shapes for Computer Vision*. arXiv preprint arXiv:2308.16139
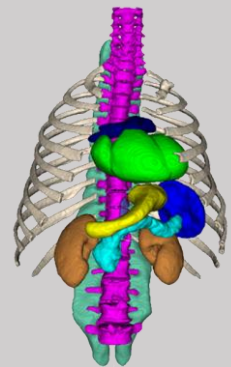
Conclusion

- Derived a benchmark dataset from whole-body segmentations for anatomical shape reconstruction (a subset of MedShapeNet)
- Proposed and evaluated a simple shape completion framework to generate whole-body pseudo-labels from partial/sparse manual annotations
- Achieved reasonable quantatitive and qualitative results

Future Work

- Include more (ideally whole-body) anatomies for multi-class anatomy completion
- Perform quantitative evaluation for each anatomy included
- Evaluate the multi-class completion framework in a whole-body segmentation task

Anatomy Completor: A Multi-class Completion Framework for 3D Anatomy Reconstruction

Jianning Li[1], Antonio Pepe[2], Gijs Luijten[1,2], Christina Gsaxner[2], Jens Kleesiek[1], Jan Egger[1]

[1]Institute for Artificial Intelligence in Medicine (IKIM), University Hospital Essen (AöR), Essen, Germany
[2]Institute of Computer Graphics and Vision (ICG), Graz University of Technology, Graz, Austria

contact: jianning.li@uk-essen.de