

Improving CT-Based X-ray Image Simulation Quality

Computer Integrated Surgery II

Spring 2022

Sean Darcy, Qiyuan Wu and Zhiyuan Ding,

under the auspices of Dr. Mathias Unberath and Ph.D candidate Benjamin Killeen

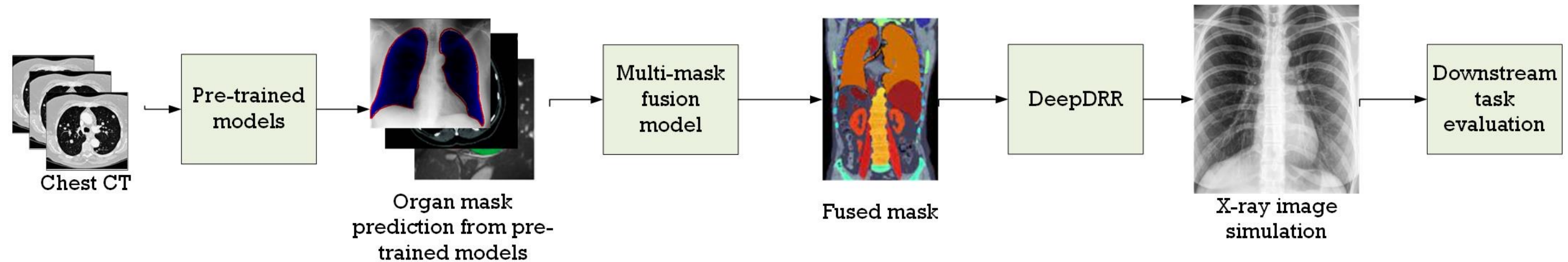


Fig 1. General project framework

Introduction

- Large X-ray image datasets are needed to effectively train artificial intelligence (A.I.) to automate fluoroscopy-guided procedures. However, expert-annotated clinical data is scarce and current simulated data is not realistic. The work in this project focused on improving the realistic quality of simulated X-Ray images.
- A multi-organ segmentation pipeline is developed for high resolution abdominal CT and integrated into DeepDRR- an X-ray simulation package.
- Digitally Reconstructed Radiograph (DRR) simulation is carried out with new segmentation pipeline, and improvements are observed with the novel pipeline developed.
- A downstream nodule detection task is proposed and implemented to compare performance of A.I. trained on either old or updated simulated datasets.

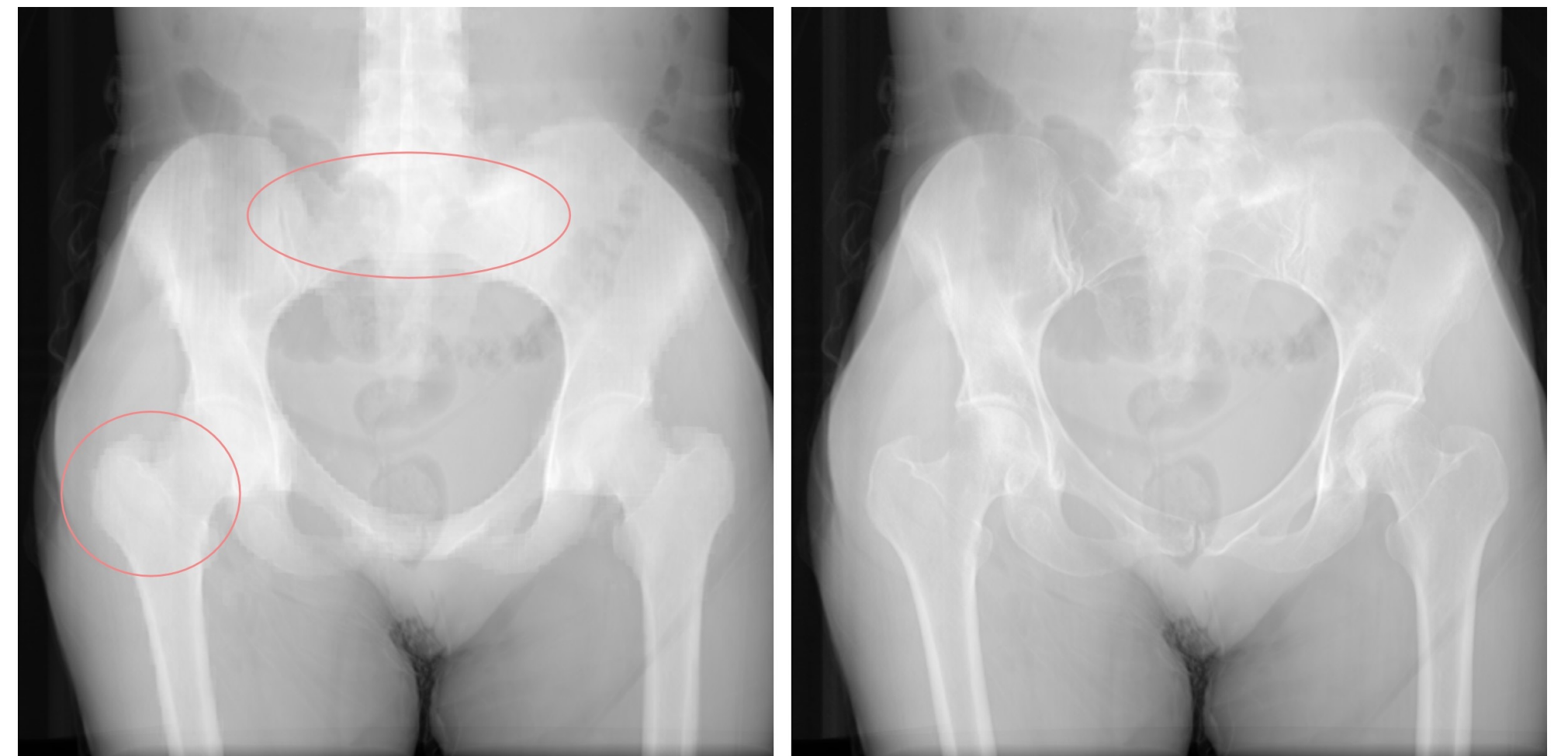


Fig 2. X-ray image simulation result comparison between the updated 3D segmentation-based approach implemented in this project (left) and the previous thresholding-based approach (right).

Outcomes and Results

- Simulation results rendered with the new pipeline, as in Fig. 1, illustrate that more anatomical details are preserved. The contrast of bone is significantly improved.
- Downstream task results are still far from what is required for deployment in a clinical setting. The DICE score for the old DeepDRR-trained model was 0.0148, while that of the new DeepDRR-trained model was 0.0315. Still, a higher test DICE score- which measures the overlap between segmentation masks and model output- for the model trained on data from our updated compared to the former version of DeepDRR is a promising result in that models trained on the updated simulated data perform better on real data in the clinic.

Future Work

- An expanded library of organ/tissue types may be introduced for X-ray image simulation with various input.
- Further integration and robustness check of current pipeline to reduce the time and computation cost for real-time simulation.
- More downstream tasks involving various organs to continue to assess clinical applicability of the simulated datasets.

Lessons Learned

- Strategies for working with large medical images.
- Evaluation design in the absence of ground-truth.

Credits

- Zhiyuan Ding: mask generation + merging
- Qiyuan Wu: mask generation + DeepDRR integration
- Sean Darcy: mask generation + downstream task evaluation

Publications

- Planning to continue for a publication over the summer.

Support by and Acknowledgements

- NIH NIBIB R21EB028505 supports this research

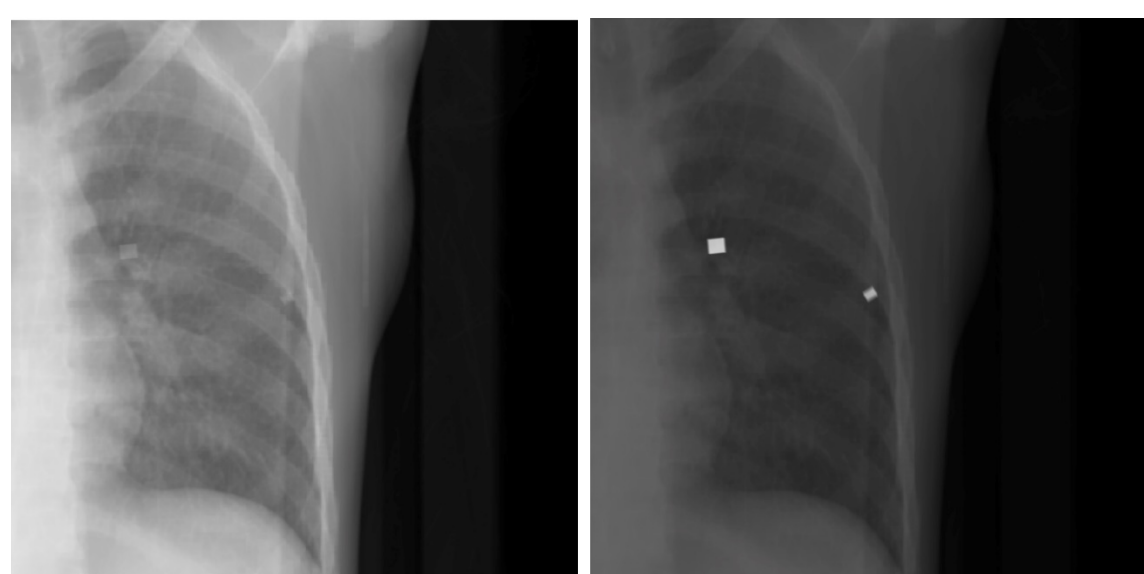


Fig 3. Lung nodule detection downstream task visualization.



Fig 4. Training and validation DICE loss curves for lung nodule detection UNets.

The Problem

- Previous DeepDRR simulation used a thresholding-based approach to assign materials coefficients to three broad tissue categories: bone, soft tissue, and air.
- These simulated images fail to preserve key anatomical X-ray image details important to clinical practice.
- An expanded simulation approach which considers the variety of organs and tissue types of the body may produce more realistic X-ray simulation results.
- No well-established quantitative method to evaluate "improved quality" or the clinical applicability of simulated data.

The Solution

- Various pre-trained deep learning tools are used to classify CT input voxels by organ/tissue type. The categories of organs/tissue type for which materials coefficients are assigned in DeepDRR is expanded from 3 (bone, soft tissue, and air) to 11 (including lung, liver, kidneys, etc...).
- The 3D-segmentation based approach to assigning materials coefficients to CT voxels is more accurate than thresholding in that key anatomical landmarks are less likely to be misclassified and a broader variety of tissue types with their own unique X-ray interaction properties are represented as they would be in real life.
- A downstream nodule detection task is proposed to assess performance of A.I. trained on either old or updated simulated datasets but tested on real X-ray data. This task serves to compare the clinical applicability of the former and updated datasets.

