Pediatric-Specific Evaluations for Deep Learning CT Image Reconstruction and Denoising

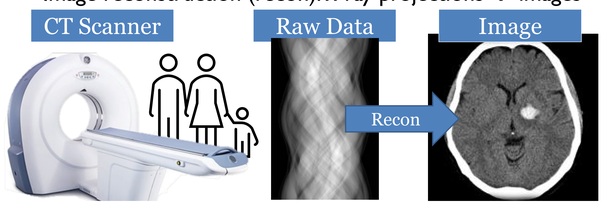
Brandon Nelson, Prabhat Kc, Andreu Badal, Shane Masters, Lu Jiang, Rongping Zeng
Division of Imaging, Diagnostics and Software Reliability | OSEL | CDRH | FDA
Brandon.Nelson@fda.hhs.gov

Introduction

Context: Pediatric CT

Image Reconstruction to Reduce Dose

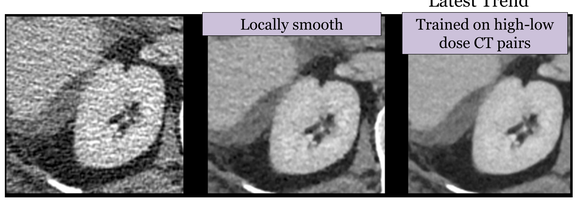
- X-ray Computed Tomography (CT) critical for pediatric care
- CT = largest contributor to pediatric radiation dose
- Image reconstruction (recon): x-ray projections → images



Deep Learning-Based Image Reconstruction (DLIR)

- Utilize deep learning to convert raw data to images
- DL denoising: DLIR that removes noise from images trained with low dose (high noise) and high dose (low noise) data

Filtered Backprojection (FBP)	Model-Based Iterative Recon (MBIR)	Deep Learning-based Image Recon (DLIR)
Conventional	Current State of the Art	Latest Trend



Advanced recons have potential to make diagnostic images with less dose

Challenge

- Deep learning recon models mainly trained on adult data
- Pediatric data scarce→efficacy of DL denoising unclear in peds

Question

- Do pediatric patients benefit equally from adult-trained DLIR models? What about subgroups: teens, toddlers, newborns?
- How to identify **performance disparities**?

Purpose: Identify Performance Disparities in Pediatric Populations

- Develop framework to assess DLIR in pediatric and adult CT:
  - in-silico image quality phantoms
  - evaluation methodology
- Demonstrate framework usage with a DL denoising model

Deep learning models designed for adults *cannot be assumed to perform equally* in pediatrics
Proposed tool evaluates performance in subgroups to identify disparities

Materials and Methods

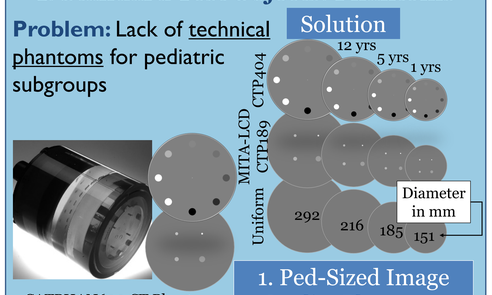
Pediatric-sized Digital Phantom Assessment Framework

Image reconstruction devices evaluated with:

1. Standard Test Objects: Phantoms

Problem: Lack of technical phantoms for pediatric subgroups

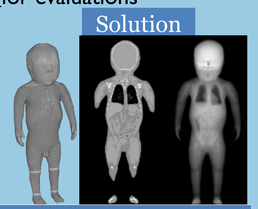
Solution



2. Clinical Assessments

Problem: Lack of pediatric patient data for evaluations

Solution



PED-ETK: PEDiatric CT Evaluation ToolKit

Test Data Generation

DL Denoising Model and Training

REDCNN model trained on 7 of the 10 **adult patient** CT exams from the Low Dose Grand Challenge Dataset

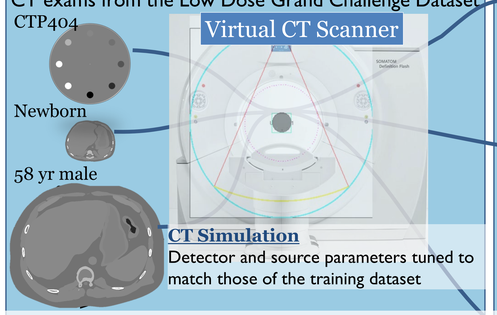
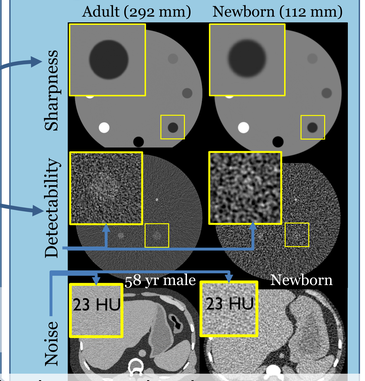


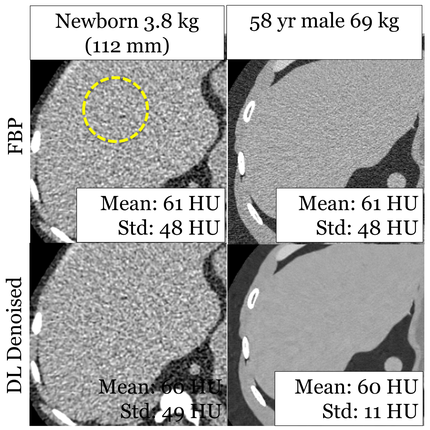
Image Quality Assessments



PED-ETK: Digital pediatric-sized image quality and anthropomorphic phantoms are virtually scanned as noisy inputs for the DLIR model followed by technical and clinically relevant task assessments

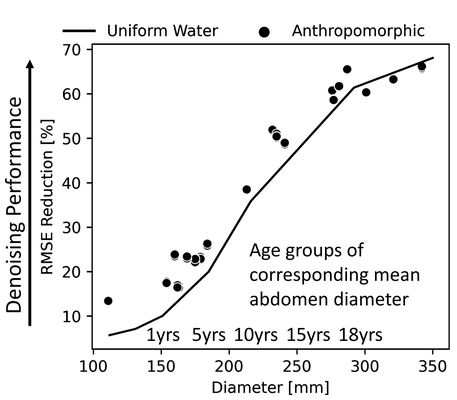
Results

Noise Reduction Assessments



Visual Assessment of Noise Reduction

- Adult and pediatric XCAT phantoms validated use of image quality phantoms to characterize DL denoising in different pediatric subgroups
- Noise reduction observed in adult phantoms
- *No visual noise change was observed in the youngest subgroups (newborns)*

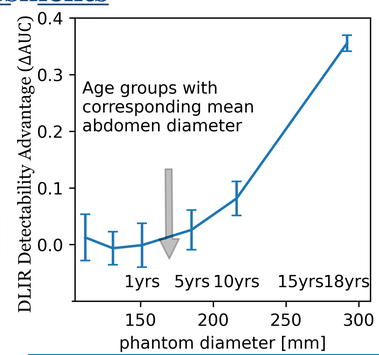


Quantitative Assessment Vs. Size (Age)

- Noise reduction: root mean square error (RMSE) reduction following DL denoising vs. diameter in uniform water and anthropomorphic phantoms
- Uniform trends with anthropomorphic at matched area, ellipse phantoms matching longest pathlength could address underestimate at high doses
- *Noise* reduction *lowest in smallest subgroups*, increasing towards adult training set size.

Clinical Task Assessments

- Medical images are made to perform clinical tasks
- Low contrast lesion detection task measured by nonprewhitening model observer
- Improvement in low contrast lesion detectability following DL denoising ***shows no advantage in youngest, smallest subgroups*** in this clinically relevant task



Discussion

- Pediatric patients are a diverse and underrepresented population
- PED-ETK provides digital phantoms of pediatric subgroups with objective and task measures of image quality to identify disparities
- Demonstrated use of PED-ETK on adult-trained REDCNN showed poor ability to generalize to pediatric patients
- Increased representation of pediatric patients in deep learning model design and evaluation is needed to ensure equal performance and benefit
- For more see: https://github.com/DIDSR/PED-ETK (available soon)

Acknowledgements

This study was supported by the National Center for Toxicological Research's **Perinatal Health Center of Excellence (PHCE)** funding program within the Food and Drug Administration