Deep Learning Reconstruction
Methods for Dosimetry and Acceleration
Jonas Teuwen
Netherlands Cancer Institute / Radboud University Medical Center

Overview
- What? Why?
- Examples
  - Digital Breast Tomosynthesis reconstruction
  - Cone Beam CT reconstruction
  - MRI reconstruction

Our rationales for deep learning based reconstruction
- Speed (neural networks have fast inference)
  - Real time adaptive MR-guided radiotherapy
    - Lower scan time
    - Faster inference (reconstruction)
- Ability to learn a data manifold
  - Dosimetry
    - Digital Breast Tomosynthesis
    - Cone-beam CT for fractionation adaptation
Reconstruction is an inverse problem

\[ y = Ax + \eta \]

**Goal:** find \( B \) such that \( x \approx By \)

We will use a neural network for this.

Data \( y \) may be strongly undersampled making inversion of \( A \) very difficult!
Deep Learning and Reconstruction

- Several approaches can be envisioned
  - Standard reconstruction algorithm + deep learning filtering
  - Directly from measured data to output
  - Model-based, combining learning with model knowledge

Learned postprocessing

Objectives

The overall objective of the Low Dose CT Grand Challenge was to quantitatively assess the diagnostic performance of developing and iterative reconstruction techniques on common low-dose patient CT datasets using a realistic task, allowing the direct comparison of the various algorithms. The main focus was on an indication of the range of performance achieved using different classes of developing or iterative reconstruction techniques.
Learned postprocessing

Fully learned approach (Automap)

LETTER

Image reconstruction by domain-transform manifold learning

Fully learned approach (Automap)
Breast Dosimetry

Digital Breast Tomosynthesis (DBT)
This information is used to reconstruct the volume.

Digital Breast Tomosynthesis (DBT)
Most deep learning reconstruction methods are *supervised*.

\[ x \approx B y \]
Most deep learning reconstruction methods are *supervised*  
\[ \mathbf{x} \approx B \mathbf{y} \]

So how to get the ground truth if this is what we have?

---

Dosimetry

- Patient BCT imaging
- Virtual compression
- Tissue classification
- Monte Carlo simulation
- 19 patients

Sechopoulos et al. (Medical Physics, 2012, 39(8), 5050-5059)
Digital Breast Tomosynthesis (DBT)

**One conclusion:** fibroglandular distribution matters for dosimetry.

Learned primal dual (Adler et al. 2018)

---

But we have this

---

Ground truth

---

DL model

---

Baseline
Average Glandular Dose

Amount of energy deposited by x-rays in glandular tissue

Amount of glandular tissue

DBT reconstruction for dosimetry

Cone-beam CT reconstruction
Goal

Ground truth

The Astra Toolbox

[Image of a medical device]

[Image of a medical scan]

[Image of a medical scan]

[Image of a medical scan]
Goal for reconstruction in adaptive radiotherapy

- Proper soft-tissue contrast to target malignant tissue
- Calibrated units for radiotherapy dose calculations
- Needs to scale to clinically relevant sizes, and have fast inference

_This would enable day-by-day adaptive radiotherapy_

**Learned SIRT**

![Diagram of Learned SIRT model](image)

Model trained on lung CTs
Generalizes well

Cone beam geometry reflecting Elekta Synergy with 60 projections
Generalizes well to real measurements

<table>
<thead>
<tr>
<th>Experiment</th>
<th>CNR</th>
<th>PRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRP (no filter)</td>
<td>1.41</td>
<td>0.13</td>
</tr>
<tr>
<td>FRP (filter)</td>
<td>6.45</td>
<td>0.17</td>
</tr>
<tr>
<td>SERT (100 iterations)</td>
<td>21.38</td>
<td>0.40</td>
</tr>
<tr>
<td>SERT (250 iterations)</td>
<td>13.97</td>
<td>0.10</td>
</tr>
<tr>
<td>SERT (low noise)</td>
<td>4.61</td>
<td>0.15</td>
</tr>
<tr>
<td>SERT (high noise)</td>
<td>21.46</td>
<td>0.25</td>
</tr>
<tr>
<td>SERT (high noise model)</td>
<td>34.32</td>
<td>0.22</td>
</tr>
<tr>
<td>SERT (patient model)</td>
<td>17.41</td>
<td>0.15</td>
</tr>
</tbody>
</table>

(a) FRP no filter  
(b) FRP with filter  
(c) SERT 250 iterations  
(d) SERT, low noise  
(e) SERT, model  
(f) High noise  
(g) SERT, patient model

Generalizes well to real measurements

(a) SERT 250 iterations  
(b) ISERT, high noise  
(c) ISERT, patient model
MRI reconstruction
The highest sampled Frequency (Bandwidth) Determines the Image Resolution

High bandwidth Low bandwidth

MRI Process
MRI samples acquired in the spatial frequency domain, aka. k-space

Proper image is retrieved by sampling at the Nyquist-rate, i.e. "fully sampled"

Scanning time is reduced by acquiring partial measurements

Corruption process given by the forward model:

Same story, use deep learning.
How to speed up? Different sampling, different artefacts

![Images from: https://people.eecs.berkeley.edu/~mlustig/CS/CSMRI.pdf]

Recurrent inference machines

- Combine convolutional layers and GRU cells
- Maintain two internal states
- And T external states (reconstructions)

MSE averaged over all T external states used for training:

\[
L(x,y) = \frac{1}{nT} \sum_{t=1}^{T} \| x_t - y_t \|^2
\]

Data Used

- 0.7mm coronal T2*-weighted brain images from 3T scanner
- 1.0mm transversal T1-weighted brain images from 3T scanner
- 0.6x0.5mm T2-weighted knee images from 3T scanner at all angles
  [http://mridata.org/fullysampled/knees](http://mridata.org/fullysampled/knees)
Model trained on all data, and cross validated

For acceleration factors 2x, 3x, 4x, and 5x on Gaussian distribution

SSIM, NRMSE, and PSNR were used as metrics of reconstruction quality

Compared with

- U-net postprocessing
- Compressed Sensing

Hyun, arXiv 2017
Lustig, Signal Processing 2008

RIMs Outperform U-nets for all accelerations

5x Accelerated T2w Knee Reconstruction

<table>
<thead>
<tr>
<th>Target</th>
<th>RIM</th>
<th>U-net</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained on: 1.0mm T1-w brain</td>
<td>0.7mm T2w brain</td>
<td>0.6/0.5mm T2w knee</td>
</tr>
</tbody>
</table>
You do not need a lot of data

Insensitivity to acceleration factor

The RIM and U-net were trained with masks at 4x acceleration
And tested on random acceleration factors sampled from U(1.5, 5.2)

Ratings by neuroradiologist

Likert scale:
5: Excellent
4: Very Good
3: Good
2: Fair
1: Poor
Questions?