Deep Learning for Biomedical Image Reconstruction: A Survey

Yuyang Xue

Sep 2021

Better health, better futures

THE UNIVERSITY of EDINBURGH | usher institute

Background

- Reconstruction is a challenging task given the ill-posed of the problem and the absence of exact analytic inverse transforms in practical cases.
- DL presents a promising approach to image reconstruction with artifact reduction and reconstruction speed-up reported in recent works as part of a rapidly growing field.
- Reconstruction algorithms transform the collected signals into a 2, 3, or 4-dimensional image.

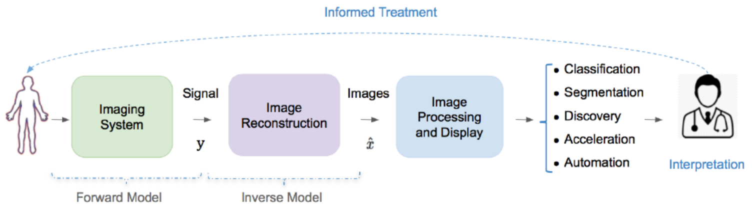

Figure: Data flow in a medical imaging and image interpretation system.

  - Forward model encodes the physics of the imaging system.
  - The inverse model transforms the collected signals by the acquisition hardware into a meaningful image.
  - The success of a diagnosis, evaluation, and treatment rely on accurate reconstruction, image visualization, and processing algorithms.

Inverse Problem and Challenges

- Let y represent a set of raw acquired sensor measurements and subject to some noise $\mathcal{N}$ intrinsic to the acquisition process. The objective is to recover the spatial-domain (or spatio-temporal) image x such that:

$$y = \mathcal{A}(\mathcal{F}(x), \mathcal{N}) \tag{1}$$

- where $\mathcal{F}(\cdot)$ is the forward model operator that models the physics of image-formation, which can include signal propagation, attenuation, scattering, reflection, and other transforms. $\mathcal{A}$ is an aggregation operation representing the interaction between noise and signal. An exact analytic inverse transform $\mathcal{A}^{-1}(\cdot)$ is not always possible.
- There may be significantly fewer measurements (M) than the number of unknowns (N).

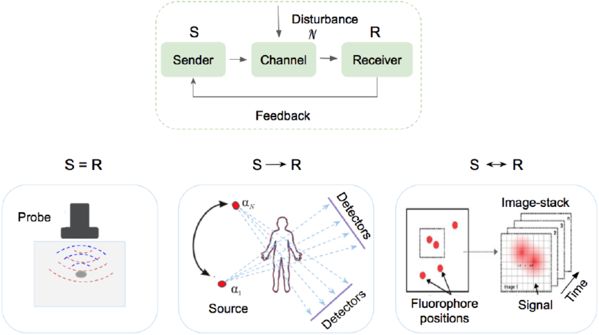

Figure: Propagation of signals from sender to receiver.

Conventional Image Reconstruction Approaches

- Analytical Methods
  - Analytical methods are based on a continuous representation of the reconstruction problem and use simple mathematical models for the imaging system.
  - Computational-expensive and can generate good image quality in the absence of noise and under the assumption of full sampling projection.
  - Ignoring the system physics and measurement noise.
  - The quality of the produced image is compromised when the signal becomes weaker.
  - Examples: The inverse Fourier Transformation for MRI.

Conventional Image Reconstruction Approaches

- Iterative Methods
  - The combine the statistical properties of data in the sensor domain (raw measurement data), prior information in the image domain, and sometimes parameters of the imaging system into their objective function.
  - Compared to analytical methods, iterative reconstruction algorithms offer a more flexible reconstruction framework and better robustness to noise and incomplete data representation problems at the cost of increased computation.
  - Iterative reconstruction methods involve minimizing an objective function that usually consists of a data term and a regularization terms imposing some prior:

$$\hat{x}^* = \arg\min_{\hat{x}} \|\mathcal{F}(\hat{x}) - y\| + \lambda \mathcal{R}(\hat{x}) \tag{2}$$

  - where $\|\mathcal{F}(\hat{x}) - y\|$ is a data fidelity term that measures the consistency of the approximate solution to the measured signal y, $\mathcal{R}(\cdot)$ is a regularization term encoding the prior information about the data, and λ is a hyper-parameter that controls the contribution of the regularization term.

Conventional Image Reconstruction Approaches

- Weakness of Iterative Methods
  - Iterative reconstruction techniques tend to be vendor-specific since the details of the scanner geometry and correction steps are not always available to users and other vendors.
  - There are substantial computational overhead costs associated with popular iterative reconstruction techniques due to the load of the projection and back-projection operations required at each iteration.
  - The reconstruction quality is highly dependent on the regularization function form and the related hyper-parameters settings ars they are problem-specific and require non-trivial manual adjustments.

Deep Learning Based Image Reconstruction

As a deep neural network represents a complex mapping, it can detect and leverage features in the input space and build increasingly abstract representations useful for the end-goal.

- Some methods considered machine learning as a reconstruction step by combining a traditional model with deep modeling to recover the missing details in the input signal or enhance the resulting image.
- A more elegant solution to reconstruct an image from its equivalent initial measurements directly by learning all the parameters of a deep nerual network, in an end-to-end fashion
- Solving for the target task directly

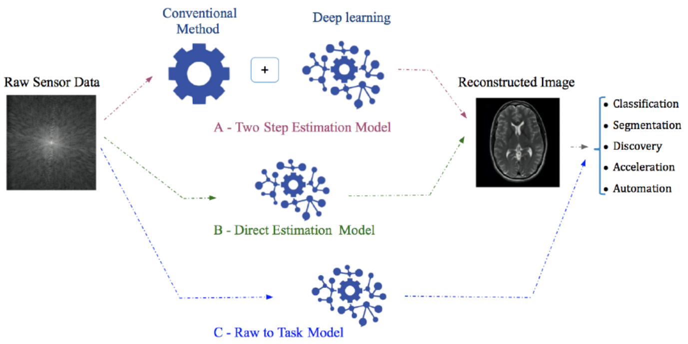

Figure: Visualization of common deep learning-based reconstruction paradigms from raw sensor data.

Deep Learning Based Image Reconstruction

- A Pre-Processing Step (Sensor Domain)
  - The problem is formulated as a regression in the sensor domain from incomplete data representation to complete data using DL methods and has led to enhanced results.
  - Examples:
    - Using U-Net to recover the unsampled k-space data folloed by an IFT to obtain the final outputimage.
    - A CT angular resolution recovery based on CNNs for accurate full view estimation from unmeasured views while reconstructing images using filtered back projection.
  - Arguments:
    - DL-based methods can fail to generalize to new test instances given the limited training dataset and DL's vulnerability to small pertubartions especially in noisy and incomplete datacase.
    - The features extracted by DL methods are limited to the sensor domain only while analytical methods' weaknesses are still present in afterword processing.
- A Post-Processing Step (Image Domain)
  - The regression task is to learn the mapping between the low-quality reconstructed image and its high-quality counterpart.
  - The main diffculty arises from the non-stationary nature of the noise and serious streaking artifacts due to information loss.
  - Examples:
    - An autoencoder is used to improve FBP results on a limited angle CT projection.
    - An adversarial network to estimate full-dose CT images from low-dose CT ones and showed it improves the model's ability to generate images with reduced aliasing artifacts.
    - De-aliasing GAN for compressed sensing MRI reconstruction that resulted in reduced aliasing artifacts while preserving texture and edges in the reconstructions.
  - Arguments:
    - The features extracted by DL methods are highly impacted by the results of the conventional methods.
    - the information missing from the initial reconstruction is challenging to be reliably recovered by post-processing like many inverse problems in the computer vision literature such as image inpainting.

Deep Learning Based Image Reconstruction

- End-to-End Image Reconstruction: Direct Estimation
  - An end-to-end solution leverages the image reconstruction task directly from sensor-domain data using a deep neural network by mapping sensor measurements to image domain while approximating the underlying physics of the inverse problem.
  - Given pairs of measurement vectors y and their corresponding ground truth images x, the goal is to optimize the parameters θ of a neural network in an end-to-end manner to learn the mapping between the measurement vector y and its reconstruction tomographic image x, which recovers the parameters of the underlying imaged tissue. We seek the inverse function $\mathcal{A}^{-1}(\cdot)$ that solves:

$$\theta^* = \arg\min_{\theta} \mathcal{L}(\mathcal{A}^{-1}(y, \theta), x) + \lambda \mathcal{R}(\mathcal{A}^{-1}(y, \theta)) \tag{3}$$

  - where $\mathcal{L}$ is the loss function penalizes the dissimilarity between the estimated reconstruction and the ground truth. The regularization term $\mathcal{R}$ if often introduced to prevent over-fitting with $\mathcal{L}1$ or $\mathcal{L}2$.

Deep Learning Based Image Reconstruction

- End-to-End: Generic Models
  - CNNs remain the most popular generic reconstruction models mainly due to their data filtering and features extraction capabilities. Specifically, encoder-decoder, U-Net, ResNet, and decoder are the most dominant architectures as they rely on a large number of stacked layers to enrich the level of feature extraction.
  - Examples:
    - WGAN based architecture for improving the image quality for 2D CT image slice reconstruction from a limited number of projection image using a combination of $\mathcal{L}1$ and adversarial losses.
    - Recurrent convolutional neural network to reconstruct high quality dynamic cardiac MR images while automatically detecting and correcting motion-related artifacts and exploiting the temoporal dependencies within the sequences.
    - A supervised sensor data distribution adaptation based MLP to take advantage of cross-domain learning and reported accuracy enhancement in detecting tissue abnormalities.
    - a supervised sensor data distribution adaptation based MLP to take advantage of cross-domain learning and reported accuracy enhancement in detecting tissue abnormalities.
    - An unsupervised CT sinograms adaptation, based on CycleGAN and content consistent regularization, to further alleviate the need for real measurement-reconstruction pairs.
  - DL based method inherently remains a black-box that can be hard to interpret. Interpretability is key not only for trust and accountability in a medical setting but also to correct and improve the DL model.

Deep Learning Based Image Reconstruction

- End-to-End: Unrolling Iterative Methods
  - Unrolling conventional optimization algorithms into a DL model has been suggested by several works in order to combine the benefits of traditional iterative methods and the expressive power of deep models.
  - Examples:
    - Deep ADMM-Net was the first proposed model reformulating the iterative reconstruction ADMM (alternating direction method of multipliers) algorithm into a deep network for accelerating MRI reconstruction, where each stage of the architecture corresponds to an iteration in the ADMM algorithm.
    - The PD-Net adopted neural networks to approximate the proximal operator by unrolling the primal-dual hybrid gradient algorithm and demonstrated performance boost compared with FBP and handcrafted reconstruction models.
    - A cascade convolutional network that embeds the structure of the dictionary learning-based method while allowing end-to-end parameter learning.
    - Qin et al. (2018) proposed to propagate learnt representations across both iterations and time. Bidirectional recurrent connections over optimization iterations are used to share and propagate learned representations across all stages of the reconstruction process and learn the spatiotemporal dependencies.

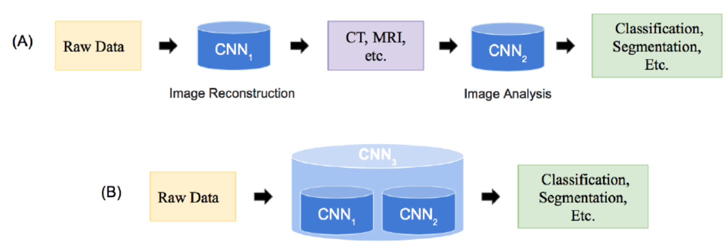

Figure: Biomedical image processing work ow usually involves two steps optimized independently (reconstruction and image analysis) for diagnosis purposes; Jointly solving these tasks using a unified model allows joint parameters tuning and feature sharing.

Deep Learning Based Image Reconstruction

- Raw-to-task Methods
  - In the raw-to-task paradigm, task results are directly inferred from raw data, where image reconstruction and processing are lumped together and reconstructed image may not be necessarily outputted.
  - Jointly solving for different tasks using a unified model is frequently considered in the computer vision field. For instance, a unified framework allows joint parameters tuning for both tasks and feature sharing where the two problems regularize each other when considered jointly.
  - Solving ill-posed problems in sequence can be more efficient than in isolation. Since the output space of a method solving an inverse problem is constrained by forming the input space of the next method, the overall solution space contracts.
  - Computational resources can then be focused on this more limited space improving runtime, solution quality, or both.

Reconstruction Evaluation Metrics

- Selected MRI datasets

<table>
<tr><th>Ref.</th><th>Year</th><th>Samp.</th><th>TA</th><th>D</th><th>Site</th><th>Arch.</th><th>E2E</th><th>Loss</th><th>Reg.</th><th>Inp.</th><th>Out.</th><th>Aug.</th><th>Metrics
M|P|S|D|C|IS</th><th>Data</th><th>Size</th><th>Pub.
Data</th></tr>
<tr><td>Zhu et al.</td><td>2018</td><td>SS</td><td></td><td>2D</td><td>Brain</td><td>Decoder</td><td>√</td><td>MSE</td><td>L2</td><td>k-space</td><td>Img</td><td>A</td><td>√ | √ |-|-|-|-</td><td>ImagNet
Clini</td><td>~10K</td><td>MGH-USCHCP</td></tr>
<tr><td>Hyun et al.</td><td>2018</td><td>SS</td><td></td><td>2D</td><td>Brain</td><td>U-Net</td><td>Pre</td><td>MSE</td><td></td><td>k-space</td><td>Full
k-space</td><td></td><td>√ |-| √ |-|-|-</td><td>Clini</td><td>1400</td><td></td></tr>
<tr><td>Cheng et al.</td><td>2018</td><td>SS</td><td></td><td>2D</td><td>Soft</td><td>CNN</td><td>Pre</td><td>MSE</td><td></td><td>k-space</td><td>Img</td><td></td><td>√ | √ | √ |-|-| √</td><td>Clini</td><td>65K</td><td></td></tr>
<tr><td>Yang et al.</td><td>2017</td><td>SS</td><td></td><td>2D</td><td>Brain</td><td>cGAN</td><td>Pos</td><td>Adver
Perc</td><td></td><td>k-space</td><td>Img</td><td>A+E</td><td>√ | √ |-|-|-| √</td><td>Clini</td><td>21K</td><td>MICCAI2013</td></tr>
<tr><td>Qin et al.</td><td>2019</td><td>SS</td><td>√</td><td>3D+t</td><td>Cardiac</td><td>RCNN</td><td>√</td><td>MSE</td><td></td><td>k-space</td><td>Img</td><td>A+E+R</td><td>√ | √ | √ |-|-| √</td><td>Sim
Clini</td><td>NA</td><td></td></tr>
<tr><td>Sun et al.</td><td>2016</td><td>SS</td><td></td><td>2D</td><td>Brain
Chest</td><td>Unrolling
ADMM</td><td>√</td><td>NMSE</td><td></td><td>k-space</td><td>Img</td><td></td><td>√ | √ |-|-|-| √</td><td>Clini</td><td>150</td><td>MICCAI2013</td></tr>
<tr><td>Schlemper et al.</td><td>2018</td><td>SS</td><td>√</td><td>2D</td><td>Cardiac</td><td>CNNs</td><td>√</td><td>MSE</td><td></td><td>k-space</td><td>Img</td><td>A+E</td><td>√ | √ |-|-|-| √</td><td>Sim</td><td>300</td><td></td></tr>
<tr><td>Oksuz et al.</td><td>2019</td><td>SS</td><td>√</td><td>2D+t</td><td>Cardiac</td><td>3D CNN
RCNN</td><td>√</td><td>MSE
CE</td><td></td><td>k-space</td><td>Img</td><td></td><td>√ | √ | √ |-|-|-</td><td>Sim
Clini</td><td>300</td><td>mridata.org</td></tr>
</table>

Figure: MRI datasets

Reconstruction Evaluation Metrics

- Quality
  - RMSE, normalized mean squared error
  - SSIM, or MS-SSIM
  - PSNR
- Inference Speed
  - Throughput measures how many problem instances can be solved over a time period
  - latency measures the time needed to process a single problem instance

Discussion

- Learning
  - DL-based approaches do not require prior knowledge, their performance can improve with it.
  - DL-based methods are more decoupled from a specific imaging (sub)modality and thus can be more generalizable.
- Interpretability
  - It is weak for the DL-based methods despite the effort in explaining the operation of DL-based methods on many imaging processing tasks.
  - Fully understanding DL-based approaches may never become practical
- Complexity
  - Conventional methods can be straightforward to implement, albeit not necessarily to design; they are often dependent on parameters requiring manual intervention for optimal results.
  - DL-based approaches can be challenging to train with a large if not intractable hyper-parameter space.
- Robustness
  - Conventional methods can provide good reconstruction quality when the measured signal is complete and the noise level is low, their results are consistent across datasets and degrade as the data representation and/or the signal-to-noise ratio is reduced by showing noise or artifacts.
  - A slight change in the imaging parameters can severely reduce the DL-based approaches' performances and might lead to the appearance of structures that are not supported by the measurements.

Future Directions

- Aim: To produce higher quality images given the most limited resource budget.
- The applicability of deep learning models to these problems is yet to be fully explored. Proposed deep learning architectures are often generic and are not fully optimized for specific applications.
- Joint multi-modal image reconstruction has been proposed on several iterative approaches to take advantage of both imaging modalities and led to improving the overall imaging performance, especially avoiding the spatio-temporal artifacts due to the scanning with different devices at different times and positions.
- Attention driven sampling received increasing interest recently especially in a limited data representation context.
- Network pruning and sparsifying present a promising direction yet to be explored on image reconstruction tasks in order to allow DL-based model processing on CPU and mobile devices.
- Reducing the encoding size in bits of network weights without altering the quality of the prediction has been demonstrated on several references deep learning image processing networks.
- Performance evaluation metrics tend to be insuffcient in the biomedical imaging context as they do not provide real diagnostic accuracy significance.