Bioengineering Department
GRAINGER COLLEGE OF ENGINEERING

Deep learning for task-based image quality assessment in radiation therapy

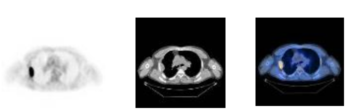
Hua Li, Ph.D.
Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL
Carle Cancer Center, Carle Foundation Hospital, Urbana, IL

Outline

- Introduction to image quality (IQ) assessment
  - Physical-based measures
  - Task-based measures and numerical observers
- Learning stochastic object models (SOM)
  - Using inter- and intra- geometric models to learn anatomical SOMs
- Task-based IQ assessment in Radiation Therapy
  - Initial characterization and demonstration with simulated CT images
- Summary and future work

Medical images

- A medical image depicts a spatial or spatial-temporal representation of some object properties
  - structural or functional properties.



PET CT PET+CT

Medical imaging systems: A generic view

- A medical imaging system maps an object property to an image.
- Black box description of data-acquisition process:

Object property → **Imaging device** → Measured data

- Mathematical description:

$$g = Hf$$

Medical imaging systems: A generic view

- A medical imaging system maps an object property to an image.
- Black box description of data-acquisition process:

Object property → **Imaging device** → Measured data

- Mathematical description:

$$g = Hf$$

Medical imaging systems: A generic view

- A medical imaging system maps an object property to an image.
- Black box description of data-acquisition process:

Object property → **Imaging device** → Measured data

- Mathematical description:

$$g = Hf$$

Medical imaging systems: A generic view

- A medical imaging system maps an object property to an image.
- Black box description of data-acquisition process:

Object property → **Imaging device** → Measured data

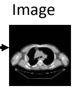
- Mathematical description:

$$g = H f$$

Computed imaging systems

- An *imaging system* refers to both the hardware and computational components:

Object property → [**Imaging device** → **Image reconstruction**] → Image

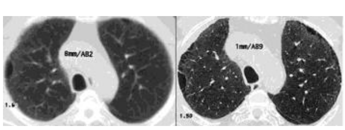


Imaging system

- Imaging systems have many tunable parameters that need to be set.

Impact of imaging parameters

- The specification of an imaging system's parameters will impact the produced image.
- Ex: Slice thickness in CT



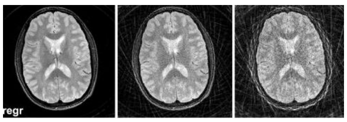
Thick 8 mm slice

Thin 1 mm slice

http://www.mevis-research.de/~hhj/Lunge/xSammlungAnaFr.html#Heart and vessels

Impact of imaging parameters: Example for MRI

- In MRI, the number of k-space samples to acquire represents a tunable parameter.
- It is desirable to minimize the number of measured samples to speed up acquisitions.



96 radials | 48 radials | 24 radials

Assessment of image quality

- In order to optimize the performance of imaging systems, figures-of-merit (FOMs) that describe image quality (IQ) are required.
- IQ metrics also permit the comparison of information contained in images acquired by different imaging modalities.
- IQ metrics can be divided into two broad classes:
  - Physical-based IQ measures
  - Task-based IQ measures

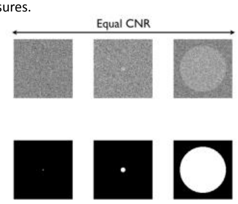
Physical-based IQ measures

- Physical measures of IQ are based on the physical and statistical characteristics of an image.
- Common measures include:
  - Spatial resolution ('sharpness of image')
  - Image contrast
  - Noise level
  - Artifact level
  - Signal-to-noise and contrast-to-noise ratios
  - RMSE (in phantom studies)

May not be easily related to the intended purpose of the image

Limitations of physical-based IQ measures: CNR

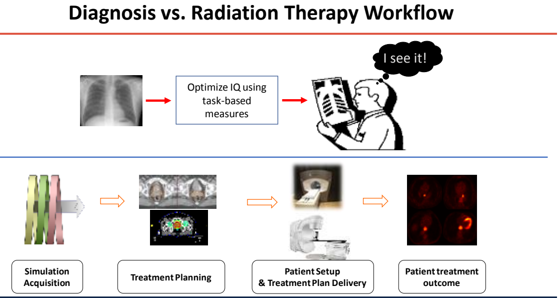
- Physical measures, such as the CNR, do not always correlate with signal detectability measures.



Task-based measures of IQ

- Task-based measures of IQ are advocated for use in evaluation and optimization of medical imaging systems
- Task-based measures of IQ quantify the ability of an observer to perform specific tasks
  - Signal detection
  - Parameter estimation
- Radiation therapy tasks:
  - Tumor/Organ-at-risk segmentation
  - RT treatment planning
  - …

Diagnosis vs. Radiation Therapy Workflow



Task-based IQ Assessment

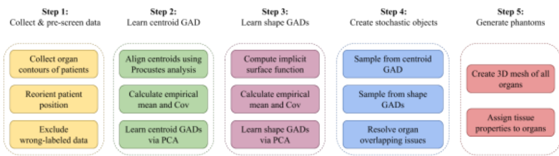
- Imaging tasks
- Observers
- Knowledge of all sources of randomness in the measured image data
- Sources of randomness in image data:
  - Measurement noise
  - Variations in the object to-be-imaged
- Stochastic object model (SOM):
  - A mathematical or computational model that describes randomness in the to-be-image object.

Learning-based SOMs for characterizing anatomical variations

Learning-based SOMs for characterizing anatomical variations

- Lack of numerical anatomical models to accurately model inter-patient and inter-organ variations in human anatomy among a broad patient population
- Available databases of high-quality volumetric images and organ contours in RT
- Development of a novel and tractable methodology for learning a SOM and generating numerical phantoms from a set of volumetric training images

Flowchart of the key steps in SOM creation process



*Steven Dolly, Yang Lou, Mark Anastasio, **Hua Li****, *"Learning-based Stochastic Object Models for Characterizing Anatomical Variations", Physics in Medicine and Biology, 2018.*

Learn the centroid geometric attributes of multiple organs

- Given *M* approved patient cases, and each contains *N* organ surfaces,
- Construct the centroid GAD by assuming that the centroids of all *N* organs follow a multi-variate Gaussian distribution:

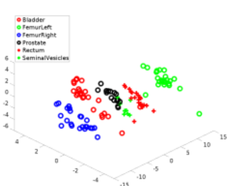
$$\mathbf{G}_m \sim \mathcal{N}(\overline{\mathbf{G}}, \mathbf{\Sigma}_{\mathbf{G}})$$

Mean: $\overline{\mathbf{G}} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{G}_m$ Covariance matrix: $\mathbf{\Sigma}_{\mathbf{G}} = \frac{1}{M-1}\sum_{m=1}^{M} (\mathbf{G}_m - \overline{\mathbf{G}})(\mathbf{G}_m - \overline{\mathbf{G}})^T$

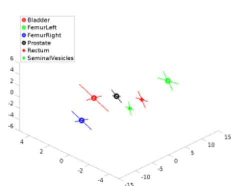
PCA analysis: $\mathbf{\Sigma}_{\mathbf{G}} = \mathbf{E}_G \mathbf{\Lambda}_G \mathbf{E}_G^T$ $\mathbf{E}_G^T(\mathbf{G}_m - \overline{\mathbf{G}}) \sim \mathcal{N}(0, \mathbf{\Lambda})$

Sampling to create random realizations: $\hat{\mathbf{G}} = \overline{\mathbf{G}} + \sum_{k=1}^{K_G} \alpha_k \sqrt{\lambda_k^G}\,\mathbf{e}_k^G$

Learned inter-structural centroid GADs



Centroid Training Set Data

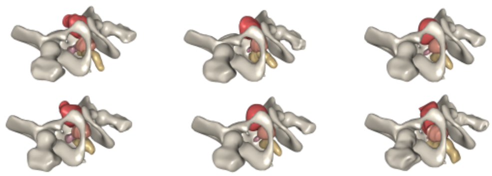


Mean Centroid and Centroid Distribution

First two principal components are visualized as line segments, whose length is the variance constrained to +/- 2δ

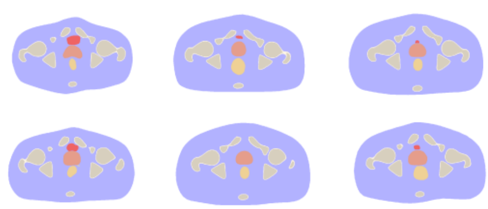
Create randomly-generated objects based on the learned GADs

- Sampling the GADs: $\hat{\mathrm{G}} = \bar{\mathrm{G}} + \sum_{k=1}^{K_G} \alpha_k \sqrt{\lambda_k^G} \mathrm{e}_k^G$ $\quad \hat{\Psi}_n = \bar{\Psi}_n + \sum_{k=1}^{K_{\Psi_n}} \beta_{k,n} \sqrt{\lambda_{k,n}^{\Psi_n}} \mathrm{e}_{k,n}^{\Psi_n}$



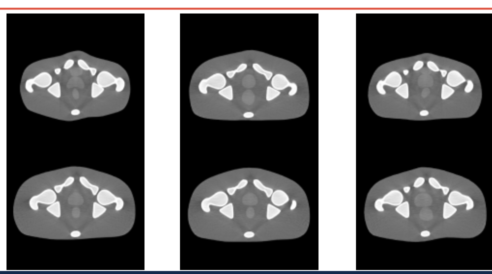
Organ models including prostate, bladder, rectum, femoral heads, pelvic bone, and seminal vesicles are displayed here, with the exception of the patient external surface for demonstration of internal organs.

Create randomly-generated objects based on the learned GADs



A sample of 2D cross-sections located at a cut plane inserted in the 3D models of the phantoms

Simulated CT images



Task-based IQ Assessment in Radiation Therapy:
Initial characterization and demonstration with simulated CT images

Task-based IQ Assessment in Radiation Therapy

- First theory developed for task-based IQA in radiation therapy based on therapeutic outcomes:

Objective assessment of image quality VI: imaging in radiation therapy

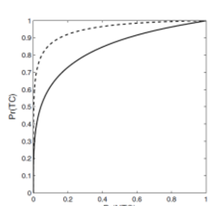
Harrison H Barrett[1,2], Matthew A Kupinski[1,2], Stefan Müeller[3], Howard J Halpern[4], John C Morris III[5] and Roisin Dwyer[6]

Phys. Med. Biol. 58 (2013) 8197–8213

- IQ Figure-of-Merit (FOM):
  - AUTOC: the area under the therapy operating characteristic (TOC) curve

Task-based IQA in Radiation Therapy

- TOC curve:
  - Plots of the probability of tumor control (TCP) vs. the probability of normal tissue complications (NTCP) as the overall dose level of a radiation treatment is varied
  - Analogy to receiver operating characteristic (ROC) curves and their variants
- TOC can defined for a single patient and also for a population of patients



Task-based IQA in Radiation Therapy

- Empirical estimation of the TCP/NTCP values from the K generated random image samples of the j-th patient:

$$\widehat{\Pr}(TC|\lambda, j) = \frac{1}{K}\sum_{k=1}^{K} \Pr(TC|\mathbf{D}(\hat{\mathbf{B}}_j^k, \mathbf{G}_j^k|\lambda), \mathbf{B}_j),$$

$$\widehat{\Pr}(NTC|\lambda, j) = \frac{1}{K}\sum_{k=1}^{K} \Pr(NTC|\mathbf{D}(\hat{\mathbf{B}}_j^k, \mathbf{G}_j^k|\lambda), \mathbf{B}_j).$$

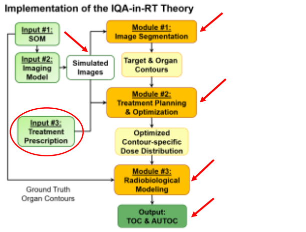
- Over an ensemble of *J* patients:

$$\widehat{\Pr}(TC|\lambda) = \frac{1}{JK}\sum_{j=1}^{J}\sum_{k=1}^{K} \Pr(TC|\mathbf{D}(\hat{\mathbf{B}}_j^k, \mathbf{G}_j^k|\lambda), \mathbf{B}_j),$$

$$\widehat{\Pr}(NTC|\lambda) = \frac{1}{JK}\sum_{j=1}^{J}\sum_{k=1}^{K} \Pr(NTC|\mathbf{D}(\hat{\mathbf{B}}_j^k, \mathbf{G}_j^k|\lambda), \mathbf{B}_j).$$

Barrett et al., Physics in Medicine and Biology, 2013.

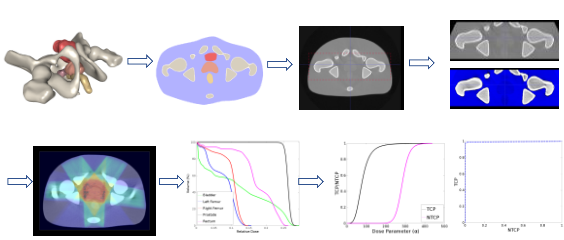
The general framework of the IQA-in-RT theory implementation



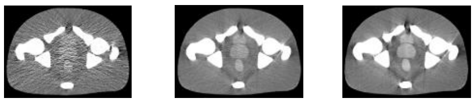
- Medical image simulation
  - CT, MRI, etc.
- Image Segmentation
  - Automatic, manual, or any other algorithms
- RT Treatment planning & optimization
  - IMRT, SBRT, etc.
- Radiobiological modeling
  - Equivalent uniform dose (EUD)
  - …
- TOC & AUTOC calculation

Steven Dolly, Yang Lou, Mark Anastasio, ***Hua Li****, *"Task-Based Image Quality Assessment in Radiation Therapy: Initial Characterization and Demonstration with Computer-Simulation Study", Physics in Medicine and Biology, 2019, in press.*

Example Implementation of the IQA-in-RT Framework

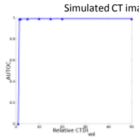


Application 1: Optimization of CT Imaging Dose



Simulated CT images with varied imaging dose

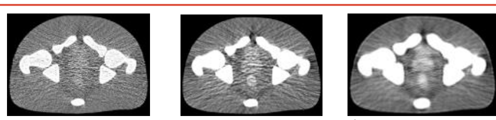
The full curve



A zoomed-in version

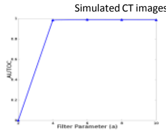
Using AUTOC as a FOM for optimizing CT imaging dose, in terms of a relative CTDI_vol

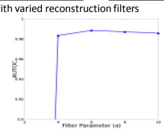
Application 2: Optimization of CT Image Reconstruction Filter



Simulated CT images with varied reconstruction filters

The full curve

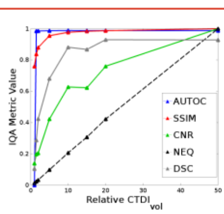




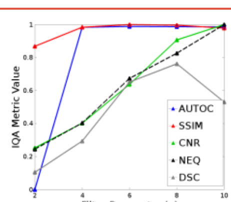
A zoomed-in version

Using AUTOC as a FOM for optimizing CT image reconstructions by optimizing the filter parameter in FBP reconstruction

Analysis 1: Comparison of AUTOC to other IQA metrics



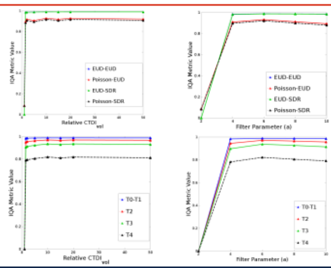
Optimizing CT imaging dose



Optimizing FBP reconstruction filter

The AUTOC compared to various IQA metrics for optimizing imaging dose and image reconstructions (the filter parameter)

Analysis 2: Sources of Bias



AUTOC vs. radiobiological model analysis of bias within the AUTOC calculation.

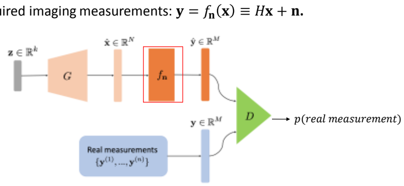
Upper: AUTOC calculations for different radiobiological models,

Lower: AUTOC calculated using different model parameters according to different prostate cancer stages.

Learning SOMs using deep generative models

AmbientGAN

- **AmbientGAN** [Bora, A., et al. ICLR. 2018]
  - Discriminator must distinguish between a *real measurement* $\mathbf{y}$ and a *simulated measurement* $\hat{\mathbf{y}}$ of the generated image $\hat{\mathbf{x}}$.
  - Acquired imaging measurements: $\mathbf{y} = f_{\mathbf{n}}(\mathbf{x}) \equiv H\mathbf{x} + \mathbf{n}$.



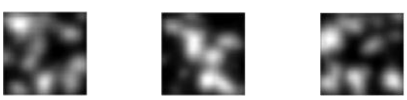
Learning SOMs using AmbientGANs

- Given a well-characterized imaging system *f* and the detector noise model *n*, AmbientGAN can be employed to learn the distribution of objects directly from noisy measurement data.
- Once trained, the generator of the AmbientGAN is the SOM.
- This represents a data-driven approach for learning SOMs from imaging measurement data.

Weimin Zhou, Sayantan Bhadra, Frank Brooks, Mark Anastasio, "Learning Stochastic Object Model from Noisy Imaing Measurements using AmbientGANs", SPIE Medical Imaging Meeting, 2019.

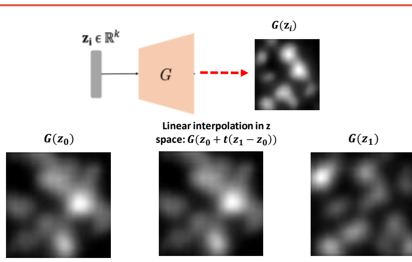
True vs. generated objects

True objects



Generated (fake) objects produced by the AmbientGAN

Walking in the latent space



Deep classifiers for task-based IQ assessment

- Consider that we have the capability to generate a set of labeled imaging measurements (e.g., via a learned SOM) or have access to a large set of labeled experimental data.
- Deep learning-based inference models (e.g., convolutional neural networks - CNNs) can be employed as numerical observers to assess signal detection-based IQ.
- Investigated supervised learning-based methodologies for approximating the Ideal Observer and Hotelling Observer test statistics for binary signal detection,
  - Employed Convolutional neural networks (CNNs) and single-layer neural networks (SLNNs) respectively.

*Weimin Zhou, **Hua Li**, Mark Anastasio, "Approximating the Ideal Observer and Hotelling Observer for binary signal detection tasks by use of supervised learning methods", IEEE Trans. on Medical Imaging, 2019.*

Summary

- Learned anatomical SOMs by characterizing the variations of human anatomy based on patient populations
- Developed the modular computational framework for implementing the task-based IQA theory in RT
- Investigated the AmbientGAN for learning SOMs from raw imaging measurements

Summary (cont'd)

- The optimization of medical imaging systems for specific diagnostic and/or RT tasks is important but challenging.
- To do so, it is important to:
  - Specify a task
  - Specify an observer
  - Account for all sources of randomness in the image data in the RT workflow
- Unsolved but critical issues:
  - Learn realistic SOMs for solving more complex RT applications
  - Build connections between computational studies and clinical practice applications
  - Approximate the performance of observers for RT applications
- Deep learning methods hold promise for addressing these problems.

Acknowledgement

- Li Group, University of Illinois at Urbana-Champaign, IL
- Anastasio Group, University of Illinois at Urbana-Champaign, IL
- Ruan Group, University of Rouen, France
- Radiation Oncology Department at Washington University in Saint Louis, MO

huali19@Illinois.edu
https://huali.bioengineering.Illinois.edu