AAPM Grand Challenges Symposium

The AAPM Working Group on Grand Challenges

S.G. Armato III

THE UNIVERSITY OF CHICAGO

Disclosures

- Royalties and licensing fees for CAD technology from the University of Chicago
- Consultant, Aduro Biotech, Inc.

Challenges

- Allows comparison of different algorithms
  - all playing by the same rules,

Challenges

- Allows comparison of different algorithms
  - all playing by the same rules,
  - on a common set of images,

Challenges

- Allows comparison of different algorithms
  - all playing by the same rules,
  - on a common set of images,
  - reporting standardized output,

Challenges

- Allows comparison of different algorithms
  - all playing by the same rules,
  - on a common set of images,
  - reporting standardized output,
  - with uniform performance assessment.

Challenges

- Controlled environment
  - database of images
  - training/testing paradigm
  - performance evaluation
  - scoring method
  - reporting structure

How to Host a Challenge

- Central organization

How to Host a Challenge

- Central organization
  - time, effort, resources

How to Host a Challenge

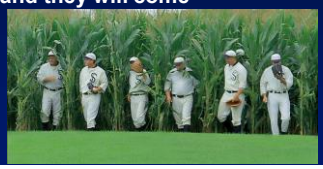
- Central organization
  - time, effort, resources, and a crystal ball

How to Host a Challenge

- Challenges will not succeed without groups willing to "accept the challenge"

How to Host a Challenge

- Challenges will not succeed without groups willing to "accept the challenge"
- "Build it and they will come"



How to Host a Challenge

- Challenges will not succeed without groups willing to "accept the challenge"
- "Build it and they will come"
- Organizers must "build it SO they will come"

How to Host a Challenge

- Challenge must be considered reputable
  - worthy of participants' time and effort
- Challenges should offer an incentive
  - conference panel or co-authorship

How to Host a Challenge

- Challenge must be considered reputable
  - worthy of participants' time and effort
- Challenges should offer an incentive
  - conference panel or co-authorship
  - fame and fortune

AAPM in the Challenge Arena

CT Low-Dose Challenge

Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 Low Dose CT Grand Challenge

Cynthia H. McCollough[a)]
Department of Radiology, Mayo Clinic, Rochester, MN 55920, USA

Adam C. Bartley and Rickey E. Carter
Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55920, USA

Baiyu Chen and Tammy A. Drees
Department of Radiology, Mayo Clinic, Rochester, MN 55920, USA

Phillip Edwards and David R. Holmes III
Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN 55920, USA

Alice E. Huang
Department of Radiology, Mayo Clinic, Rochester, MN 55920, USA

Farhana Khan
Information Services, American Association of Physicists in Medicine, Alexandria, VA 22314, USA

Shuai Leng, Kyle L. McMillan, Gregory J. Michalak, Kristina M. Nunez, Lifeng Yu, and Joel G. Fletcher
Department of Radiology, Mayo Clinic, Rochester, MN 55920, USA

(Received 28 October 2016; revised 11 May 2017; accepted for publication 11 May 2017; published 12 October 2017)

CT Low-Dose Challenge

A fast method to emulate an iterative POCS image reconstruction algorithm

Gengsheng L. Zeng[a)]
Department of Engineering, Weber State University, Ogden, UT 84408, USA
Department of Radiology and Imaging Sciences, University of Utah, Salt Lake City, UT 84108, USA

A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction

Eunhee Kang,* Junhong Min,* and Jong Chul Ye[a)]
Bio Imaging and Signal Processing Lab., Dept. of Bio and Brain Engineering, KAIST, Daejeon, Korea

Low-dose CT reconstruction using spatially encoded nonlocal penalty

Kyungsang Kim, Georges El Fakhri, and Quanzheng Li[a)]
Gordon Center for Medical Imaging, Massachusetts General Hospital and Harvard Medical School, 125 Nashua Street 6th floor, Suite 660, Boston, MA 02114, USA

SPIE-AAPM-NCI LUNGx Challenge

- Computerized lung nodule classification
  - Armato et al. JMI, 2015.
  - Armato et al. JMI, 2016.
- Collaboration among SPIE, AAPM, and NCI
  - University of Chicago, University of Michigan, and Oak Ridge National Laboratory

SPIE-AAPM-NCI PROSTATEx Challenges

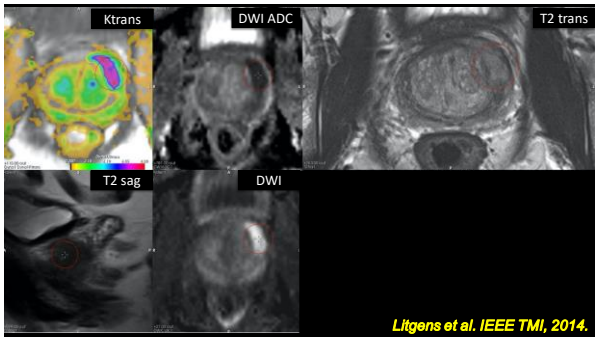
- PROSTATEx: SPIE Medical Imaging 2017
- PROSTATEx-2: AAPM Annual Meeting 2017

SPIE-AAPM-NCI PROSTATEx Challenges

- PROSTATEx: SPIE Medical Imaging 2017
- PROSTATEx-2: AAPM Annual Meeting 2017
- Same data: multi-parametric prostate MR scans
- Two diagnostic tasks

SPIE-AAPM-NCI PROSTATEx Challenges

- PROSTATEx: SPIE Medical Imaging 2017
- PROSTATEx-2: AAPM Annual Meeting 2017
- Same data: multi-parametric prostate MR scans
- Two diagnostic tasks
- Collaboration among AAPM, SPIE, and NCI
  - University of Chicago, University of Michigan, Radboud University, and FDA



SPIE-AAPM-NCI PROSTATEx Challenges

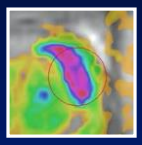
- PROSTATEx must not bias PROSTATEx-2

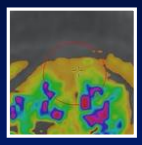
PROSTATEx

SPIE-AAPM-NCI Prostate MR Classification Challenge

PROSTATEx

Task: Distinguish between "clinically significant" and "not clinically significant" prostate lesions on MR





PROSTATEx

Task: Distinguish between "clinically significant" and "not clinically significant" prostate lesions on MR

Timeline:

- November 21, 2016: training cases available
  - 330 lesions with truth (binary clinical significance)
  - spatial location
  - zone (anterior fibromuscular stroma, peripheral, seminal vesicle, transition)

PROSTATEx

Task: Distinguish between "clinically significant" and "not clinically significant" prostate lesions on MR

Timeline:

- November 21, 2016: training cases available
  - 330 lesions with truth (binary clinical significance)
- December 12, 2016: test cases available
  - 208 lesions without truth
  - spatial location
  - zone

PROSTATEx

Task: Distinguish between "clinically significant" and "not clinically significant" prostate lesions on MR

Timeline:

- November 21, 2016: training cases available
  - 330 lesions with truth (binary clinical significance)
- December 12, 2016: test cases available
  - 208 lesions without truth
- February 15, 2017: SPIE Medical Imaging 2017

PROSTATEx

Task: Distinguish between "clinically significant" and "not clinically significant" prostate lesions on MR

- submit single real number on [0, 1] for each lesion representing likelihood of clinical significance
  - performance assessment: ROC analysis

PROSTATEx-2

SPIE-AAPM-NCI Prostate MR Gleason Grade Group Challenge

PROSTATEx-2

Task: Develop quantitative MR biomarkers for the determination of Gleason Grade Group in prostate cancer

PROSTATEx-2

Task: Develop quantitative MR biomarkers for the determination of Gleason Grade Group in prostate cancer

Timeline:

- May 15, 2017: training cases available
  - 112 lesions with truth (Gleason Grade Group)
  - spatial location
  - zone (anterior fibromuscular stroma, peripheral, seminal vesicle, transition)

PROSTATEx-2

Task: Develop quantitative MR biomarkers for the determination of Gleason Grade Group in prostate cancer

Timeline:

- May 15, 2017: training cases available
  - 112 lesions with truth (Gleason Grade Group)
- June 5, 2017: test cases available
  - 70 lesions without truth
  - spatial location
  - zone

PROSTATEx-2

Task: Develop quantitative MR biomarkers for the determination of Gleason Grade Group in prostate cancer

Timeline:

- May 15, 2017: training cases available
  - 112 lesions with truth (Gleason Grade Group)
- June 5, 2017: test cases available
  - 70 lesions without truth
- August 1, 2017: AAPM Annual Meeting 2017

PROSTATEx-2

Task: Develop quantitative MR biomarkers for the determination of Gleason Grade Group in prostate cancer

- submit ordinal value on [1, 5] for each lesion representing Gleason Grade Group
  - performance assessment: quadratic-weighted kappa

2019 Digital Pathology Challenge

SPIE-AAPM-NCI CellularityX Challenge

SPIE-AAPM-NCI CellularityX Challenge

- Develop algorithms to automatically assess cellularity in pathology whole slide image patches
- Cellularity assessed as
  - a score and a categorical classification
    - 0 (normal)
    - 1–30 (low cellularity)
    - 31–70 (medium cellularity)
    - 71–100% (high cellularity)

SPIE-AAPM-NCI CellularityX Challenge

- Reference standard: clinical assessment of cellularity by two pathologists

SPIE-AAPM-NCI CellularityX Challenge

Stay tuned for further details

AAPM Working Group on Grand Challenges

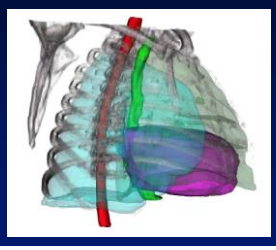
AAPM Working Group on Grand Challenges

- Mission 1
  - Identify “hot topic” challenge tasks and organize internally

AAPM Working Group on Grand Challenges

- Mission 1
  - Identify “hot topic” challenge tasks and organize internally
- Mission 2
  - Solicit challenge proposals from the community
  - Assist with resources and expertise

Auto-Segmentation for Thoracic Radiation Treatment Planning Challenge



AAPM Working Group on Grand Challenges

Proposals seek:

- title of challenge (including proposed nickname)
- background on the issue to be addressed
- statement of the unmet need to be served

AAPM Working Group on Grand Challenges

Proposals seek:

- challenge methodology and logistics
  - source and numbers of cases for training/testing
  - parameters to be held fixed or to be “matched” across cases
  - reference standard

AAPM Working Group on Grand Challenges

Proposals seek:

- challenge methodology and logistics
  - proposed challenge hosting platform
  - proposed data hosting platform
  - performance assessment metric(s)
  - publication plan
  - plan for data persistence

AAPM Working Group on Grand Challenges

Proposals seek:

- WGGC assistance requested (logistical, financial)

AAPM Working Group on Grand Challenges

Proposals seek:

- WGGC assistance requested (logistical, financial)

We want to help you “build it SO they will come”

Sparse-View Reconstruction Challenge for 4D Cone-Beam CT (SPARE Challenge)

Sparse-View Reconstruction Challenge for 4D Cone-Beam CT (SPARE Challenge)

- Participants developed algorithms to reconstruct cone-beam CT images from sparsely sampled projection data

Sparse-View Reconstruction Challenge for 4D Cone-Beam CT (SPARE Challenge)

- Participants developed algorithms to reconstruct cone-beam CT images from sparsely sampled projection data
- Organized by Andy Shieh, The University of Sydney
  - with WGGC support

Dataset

- MR scans of each patient contained
  - Transaxial and sagittal T2-weighted images
  - K_{trans} images (computed from dynamic contrast-enhanced images)
  - Apparent diffusion coefficient (ADC) images (computed from diffusion-weighted images)
- Thumbnail of each lesion for visual reference

Litjens et al. IEEE TMI, 2014.

Rules

- Train with training set without restriction
- No human intervention with test set, except for manual or human-supervised delineation of prostate boundary or gross lesion margin
- Participants not allowed to withdraw from the Challenge after test set groupings submitted

Challenges

- Friendly competition
- Educational
- Participants deserve much credit

Challenge Considerations

- Database

Challenge Considerations

- Database
  - How many cases? balance effort/cost vs. power
  - Technical parameter consistency
    - slice thickness
    - contrast enhanced
  - Lesion characteristics consistency
    - size
    - histology

Challenge Considerations

- Database
- Instructions

Challenge Considerations

- Database
- Instructions
  - Acceptable level of human input
  - Appropriate use of test set
  - Don’t assume

Challenge Considerations

- Database
- Instructions
- Clinical data

Challenge Considerations

- Database
- Instructions
- Clinical data
  - limit to DICOM header (age, gender)?

Challenge Considerations

- Database
- Instructions
- Clinical data
- Commitment by participants

Challenge Considerations

- Database
- Instructions
- Clinical data
- Commitment by participants
  - anonymity

Challenge Considerations

- Database
- Instructions
- Clinical data
- Commitment by participants
- Epilogue

Challenge Considerations

- Database
- Instructions
- Clinical data
- Commitment by participants
- Epilogue
  - image data persistence
  - publish?

Gleason Score (Grade)

- Reflects the appearance of cancer cells from biopsy
  - 1 = similar to normal prostate cells
  - 5 = very abnormal

http://urologyrbh.blogspot.com/2011/06/prostate-biopsy.html

Gleason Score (Grade)

- Reflects the appearance of cancer cells from biopsy
  - 1 = similar to normal prostate cells
  - 5 = very abnormal
- Two ratings assigned
  - one to the most dominant cell appearance
  - one to the second-most dominant appearance
- The two ratings are combined: Gleason score [2-10]

Gleason Score (Grade)

- Clinically meaningful binning of Gleason scores

Gleason Grade Group	Gleason Score
1	≤ 6
2	7 (3 + 4)
3	7 (4 + 3)
4	8
5	9 or 10

Epstein et al. Am J Surg Pathol, 2016.