Grand Challenges with MIDRC

Sam Armato, Ph.D.
The University of Chicago
“Grand Challenges”

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

• Controlled environment
  • database of images
  • training/testing paradigm
  • performance evaluation
  • scoring method
  • reporting structure
Successful Challenges

• Challenges will not succeed without groups willing to “accept the challenge”

• “Build it and they will come”
Successful Challenges

• 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”
Successful Challenges

• Challenge must be considered reputable
  • worthy of participants’ time and effort

• Challenges should offer an incentive
  • conference panel or co-authorship or prize money
Why Challenges?

• Advance the field
Why Challenges?

• Advance the field
  • motivate groups around the world to dedicate their time and talents toward a defined task
  • reveal approaches to the task that are (1) promising or (2) weaker
  • foster open science with code depositing
Challenge Quiz

- What is the single biggest challenge of challenges?
Challenge Quiz

• What is the single biggest challenge of challenges?

Image Data
Challenge Quiz

https://wvhistoryonview.org/catalog/005796
Challenge Quiz

https://wvhistoryonview.org/catalog/005796
https://www.target.com/p/toy-story-4-pig-ham-coin-bank-pink/
Challenge Quiz

- But you also want....
  - image metadata
  - clinical data
  - demographic data
  - patient/image/lesion labels
  - expert annotations
Challenges are contained within
- 3 of the 5 Technology Development Projects
- 2 of the 12 Collaborative Research Projects
AAPM Challenges

- LUNGx (2015)
- PROSTATEx / PROSTATEx-2 (2017)
- BreastPathQ (2019)
- DBTex (2020)
- Proffered challenges (6+)
RSNA Challenges

- Pediatric Bone Age Challenge (2017)
- Pneumonia Detection Challenge (2018)
- Intracranial Hemorrhage Detection Challenge (2019)
- Pulmonary Embolism (2020)
ACR Challenges

• Breast Density Crowdsourcing Challenge (2019)

• Resident Mammography Challenge (2019)

• SIIM-ACR Pneumothorax Challenge (2019)
MIDRC Challenge Process
MIDRC Challenge Process

- Most challenges develop from an existing, static, limited set of images and associated data

- Challenges are designed based on the limitations of this key ingredient

- The magnitude and expanding nature of MIDRC actually make data collection for challenges a challenge
Test data request form has been developed
MIDRC Challenge Process

- Test data request form has been developed
  - Allows for challenge organizers to
    - collect their thoughts
    - work through their challenge logistics
    - identify their challenge needs
MIDRC Challenge Process

• Test data request form has been developed
  • Challenge title
  • Number of subjects
  • Number of imaging studies
  • Distribution of COVID-19 cases and other disease
  • Reference standard requirements
  • Patient demographic distributions (age, gender, race)
  • Need for associated clinical data
  • Timeline
ACR or RSNA (de-Id and QC’d data) with images and complete* clinical, demographic, and acquisition data

**Gen3 AWS Bucket**

**Sequestered?**

- **YES**
  - Sequestered MIDRC Data Commons
- **NO**
  - **Use for challenges?**
    - **YES**
      - Use in challenge, then publish
    - **NO**
      - Public MIDRC Data Commons (data.midrc.org)

* Images and other data will **ONLY** be passed to Gen3 once all data is available
Conclusion

• MIDRC will provide an extensive resource of data for a vast array of grand challenges

• Initially, MIDRC challenges will be organized by MIDRC investigators

• Future MIDRC challenges will be organized by the medical imaging research community at large