Artificial Intelligence: Why does it matter to little people?

Sudhen B. Desai, MD, FSIR
Director of Research, Interventional Radiology
BOD, Society of Physician Entrepreneurs

DISCLOSURES

- None Relevant

BACKGROUND

- Publications on AI in radiology have increased from 100–150/y to 700–800/y over last decade
- Open data sets for training – NIH, RSNA
- Radiology: Artificial Intelligence as stand-alone journal
- RSNA hosts annual AI Challenge
  - 2018: Pneumonia Detection
  - 2017: Pediatric Bone Age

CHALLENGES

- Data is KING!
  - Free-standing children's hospitals need to have enough data to train models = patient volumes
  - No large open data sets for Pediatric Radiology AI Training
- Resources
  - FSCH are organized different than traditional Academic Medical Centers
  - Most lack strategic, interoperable and agile enterprise IT infrastructure
    - Heavy focus on HIS, PACS, EMR as functionally integrated systems but difficult to transfer huge swaths of data throughout

CHALLENGES

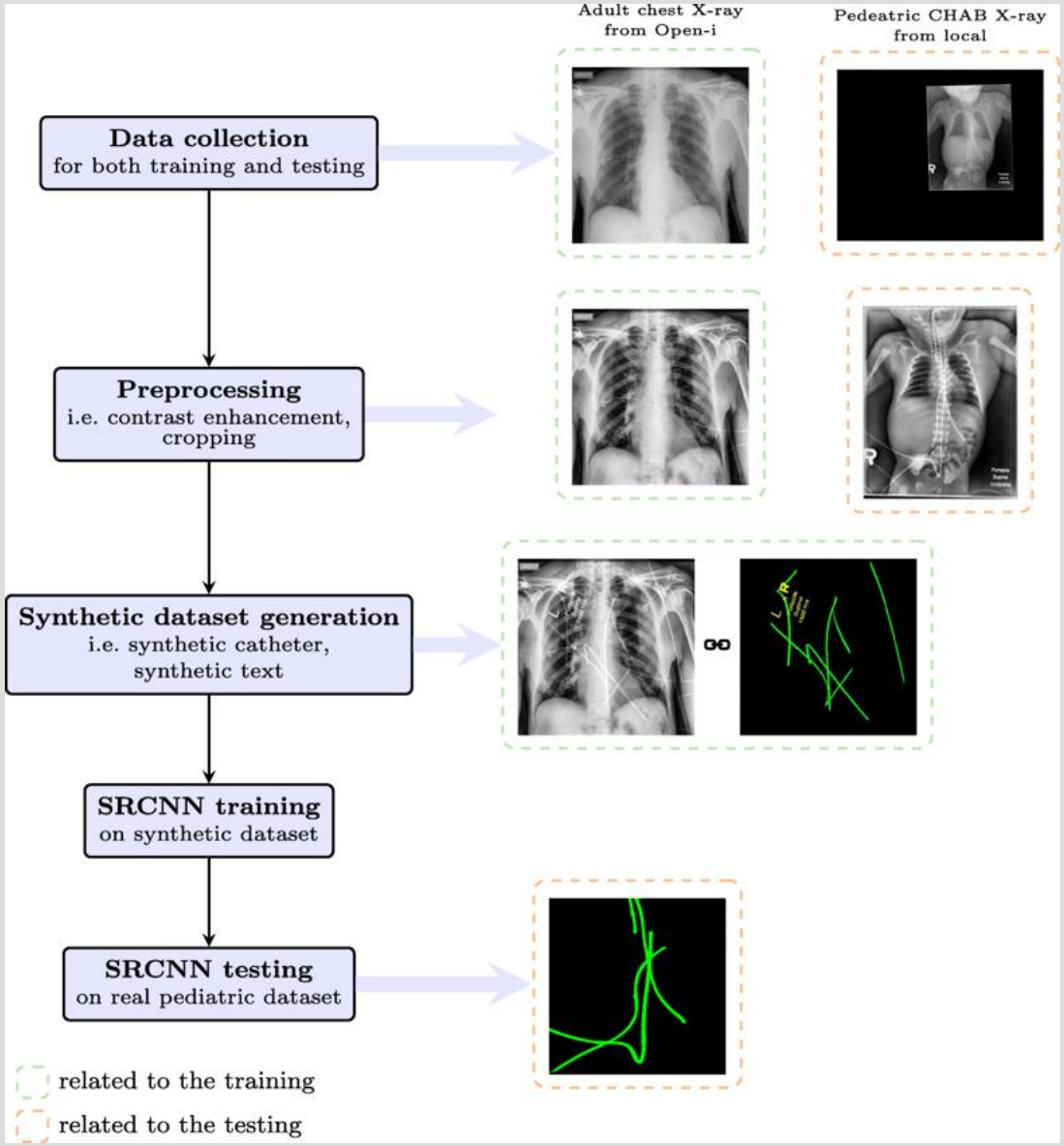
- Information Density
  - Images overall smaller – greater anatomy presented (e.g. adult CXR vs. ChAB)
- Defining “normal development” versus pathology
  - “Normal” varies per age – e.g. growth centers
  - Growth centers versus fracture fragments
  - Nutrient channels versus fractures
- Pediatric Radiologists are not Data Scientists

CATHETER DETECTION

- Every pediatric radiograph reviewed for lines/tubes
  - ETT, NGT, OGT, UVC, UAC, PICC, tCVL, CVL, IO, EPL
- Confounding issues:
  - EKG leads, casting materials, blankets, image text

ANALYSIS

- Data Collection - adult CXR from NIH open data set
- Preprocessing - contrast enhancement, cropping
- Synthetic Dataset Generation - synthetic catheter, synthetic text
- SRCNN Training
- SRCNN Testing



Texas Children's

Yi, X., Adams, S., Babyn, P. et al. J Digit Imaging (2019). https://doi.org/10.1007/s10278-019-00201-7

VALIDATION

- Capable of detecting umbilical catheters, NGTs and ETTs < 1s when deployed on a single GPU

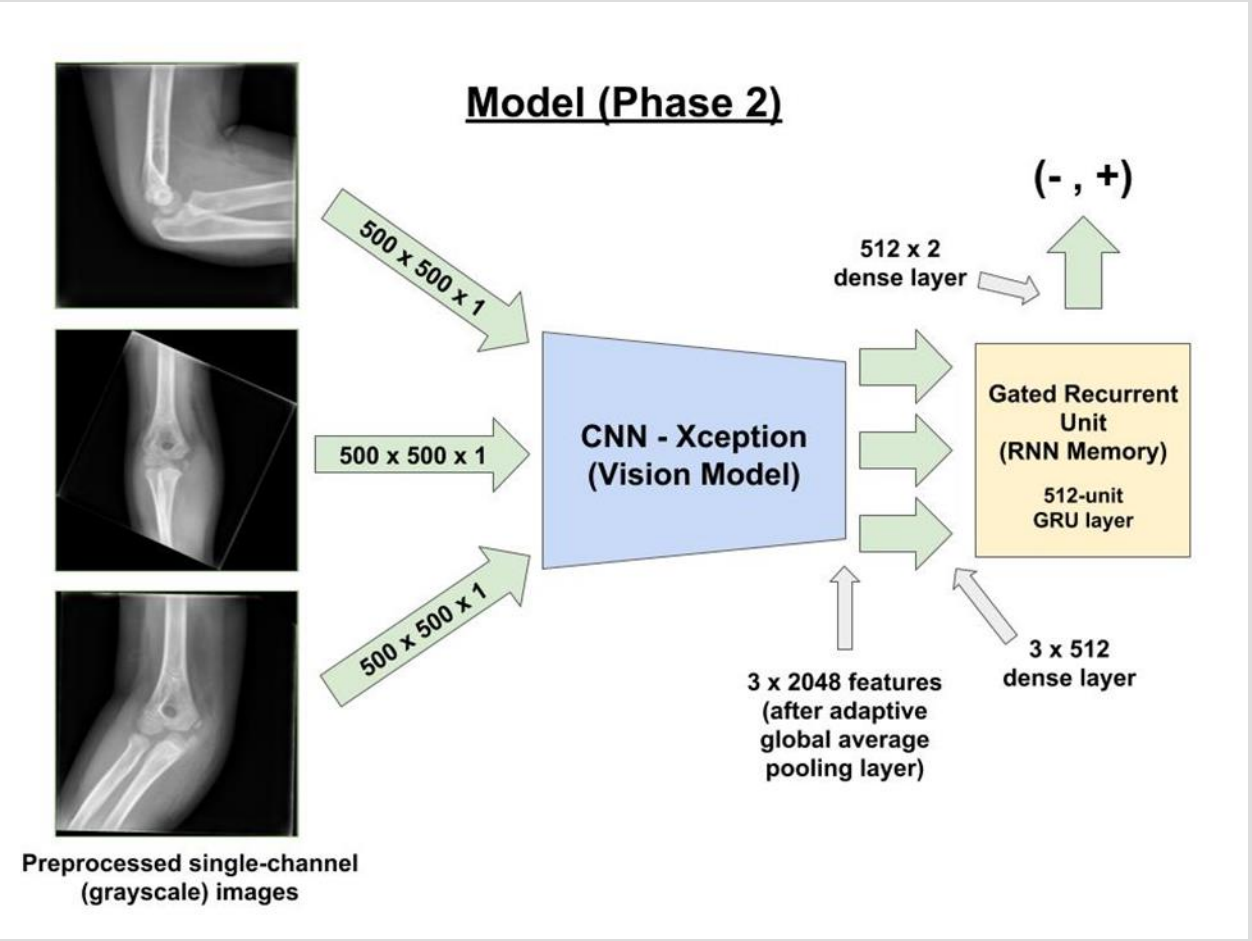
- F_{β}=0.8

BACKGROUND (TCH)

- ~700 pediatric beds through the enterprise
- 1 pediatric radiologist covers 3 ER (one level 1, two level 2) centers and 13 urgent care centers during the after-hours period
- Difficult to triage EC volumes – which film is most important?
  - 193,556 radiographic exams total in 2018 (>500/d)
    - 37,419 ordered from the EDs
  - 16,182 CT scans
  - 20,727 MR exams

ELBOW RADIOGRAPHS

- Reviewed data from > 21,000 pediatric elbow examinations (~59K images)
  - January 2014 to December 2017
  - Training set of > 20,000 studies
  - Remaining studies were used as a validation set
- Images were only interpreted for research purposes, not for clinical implementation


Model (Phase 2)
500 x 500 x 1
500 x 500 x 1
500 x 500 x 1
CNN - Xception
(Vision Model)
512 x 2
dense layer
(- , +)
Gated Recurrent
Unit
(RNN Memory)
512-unit
GRU layer
3 x 2048 features
(after adaptive
global average
pooling layer)
3 x 512
dense layer
Preprocessed single-channel
(grayscale) images

ANALYSIS

- Area under the receiver operating characteristic curve - 0.95
- Accuracy – 88%
- Sensitivity – 91%
- Specificity – 84%

VALIDATION

- Recurrent neural network could classify entire series of images instead of just making a classification based on a single image

- Begins to model the image analysis algorithms of Radiologists

LOWER EXTREMITY LONG BONES

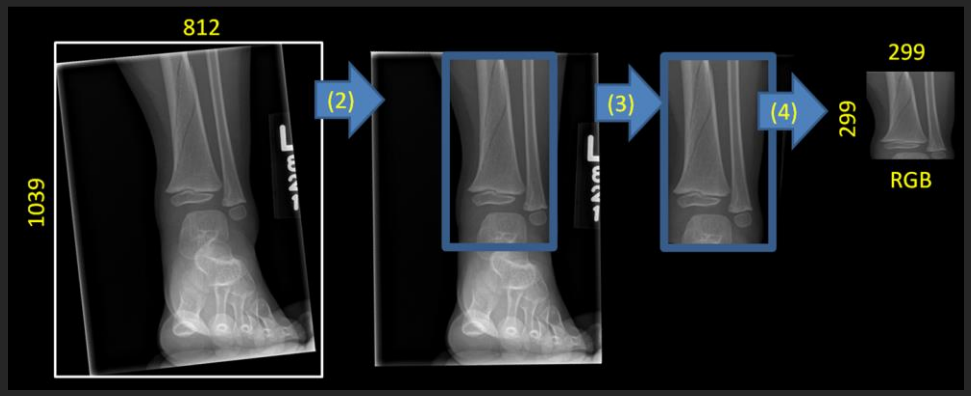
- Tibial fractures
- ** 516 ** studies of the foot, ankle, tibia and fibula (2118 radiographs)
- 2009-2017

ANALYSIS

- AUC - 0.995
- Accuracy - 97.8%
- Sensitivity - 95.9%
- Specificity - 100%

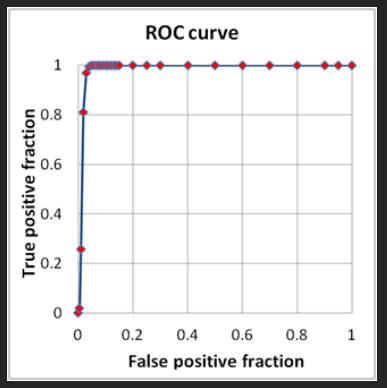
ANALYSIS

Data pre-processing

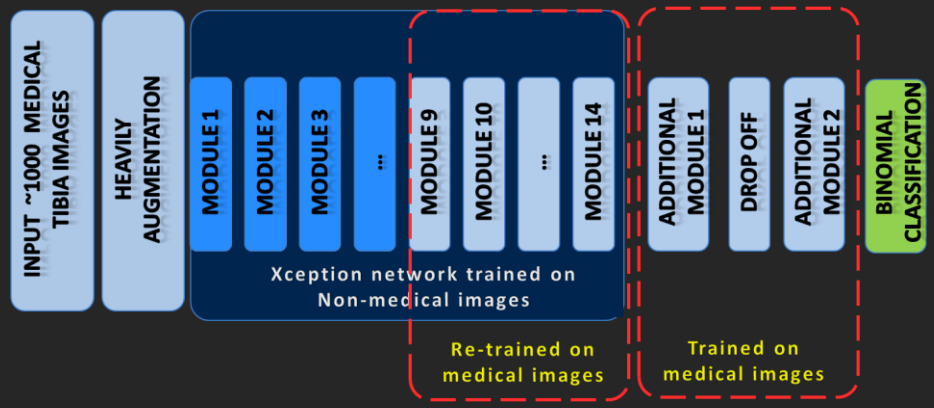


Results

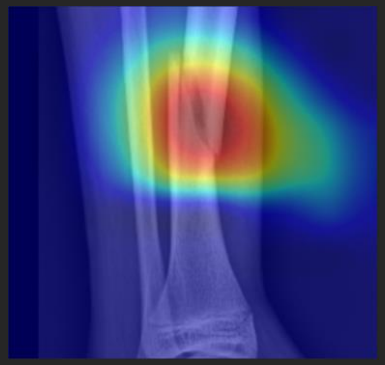
- Accuracy 97.8%
- Sensitivity 95.9%
- Specificity 100%

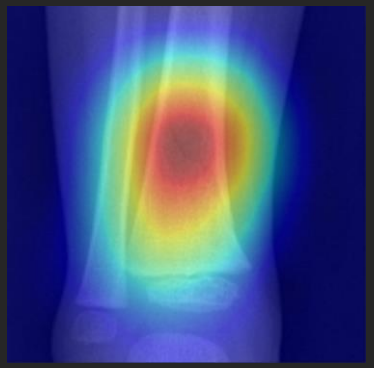


Network architecture (Transfer Learning)



Gradient-weighted Class Activation Mapping

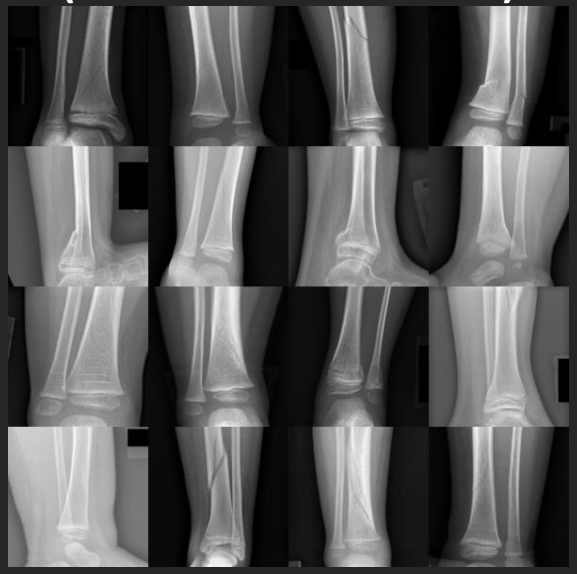




Texas Children's

ANALYSIS

CNN-DTF
(Distal Tibial Fractures)



ACC: 97.8%

CNN-LBF
(Long Bone Fractures)

ACC: 86.6%

CNN-DTF tested on LBF

ACC: 97.8%	drop to 58.5%
Sensitivity: 95.8%	drop to 92.9%
Specificity: 100%	drop to 24.2%

CNN-LBF tested on DTF

ACC: 86.6%	drop to 66.7%
Sensitivity: 95.8%	drop to 43.4%
Specificity: 100%	drop to 89.8%

The results show the importance of utilizing anatomy-specific training sets, a significant limitation of current CNN's

Texas Children's

VALIDATION

- Transfer learning allowed small training set to be utilized (~500 patient studies, ~2100 radiographs)
  - ** approximately what a trainee might see during residency and Peds Radiology fellowship

CONCLUSIONS

- AI can enhance pediatric radiologists' efficiency, proficiency, and accuracy
- Triage urgent/acute care imaging
- Transfer learning can markedly reduce training set volumes
  - Helpful to address concerns in Peds AI

FUTURE

- Improve cross-sectional imaging efficiency
  - Embedding AI algorithms may allow reduced radiation dosing and contrast dosing – IMPORTANT!!!
  - Reduce in-magnet time for MRI
    - Reduce Anesthesia time in Peds
- Allow improved inter-disciplinary coordinations in clinical care
  - NAT

ACKNOWLEDGEMENTS

- Ananth Annapragada, Ph.D.
- Zbigniew A. Starosolski, Ph.D.
- Herman Kan, MD
- Jesse Rayan, MD
- Nakul Reddy, MD

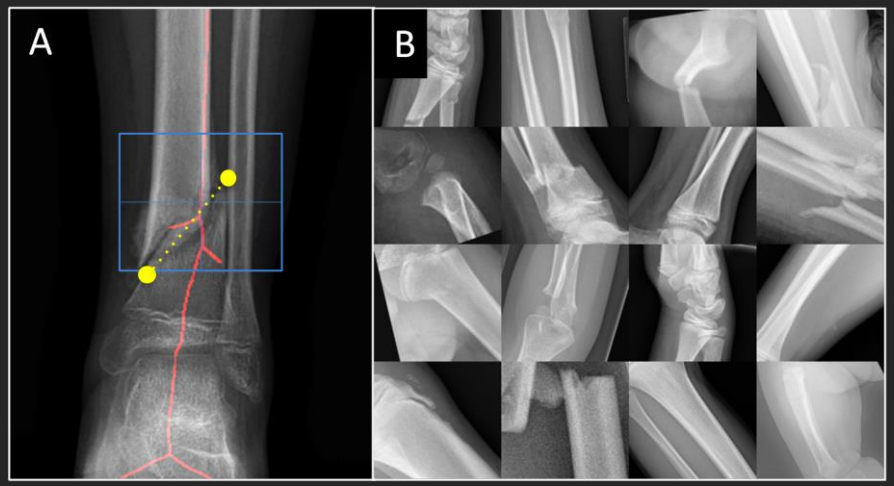
TCH PEDIATRIC RADIOLOGY AI SYMPOSIUM

- 10/28-29/2019
- Houston, TX
- sbdesai@texaschildrens.org
- Hope you can join us!

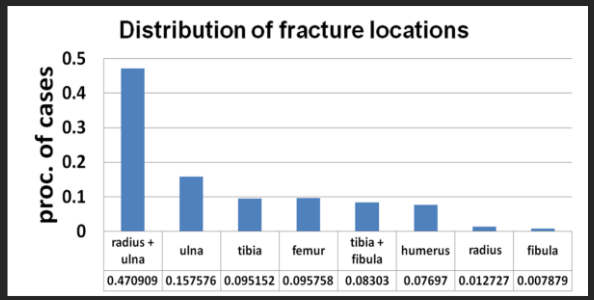
Texas Children's®

COMMENTS/QUESTIONS?

Data pre-processing



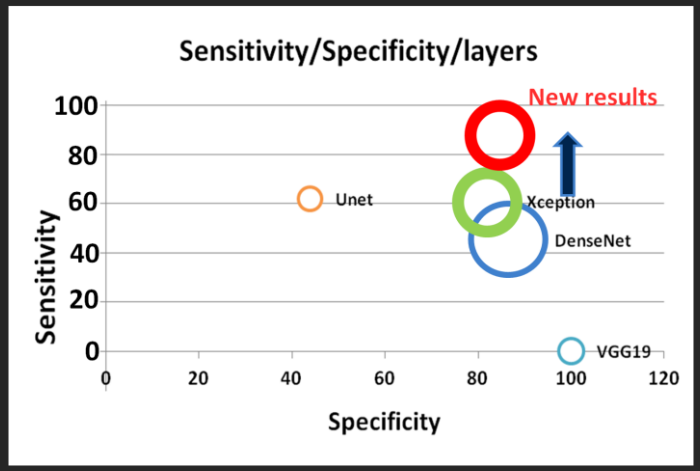
Data distribution

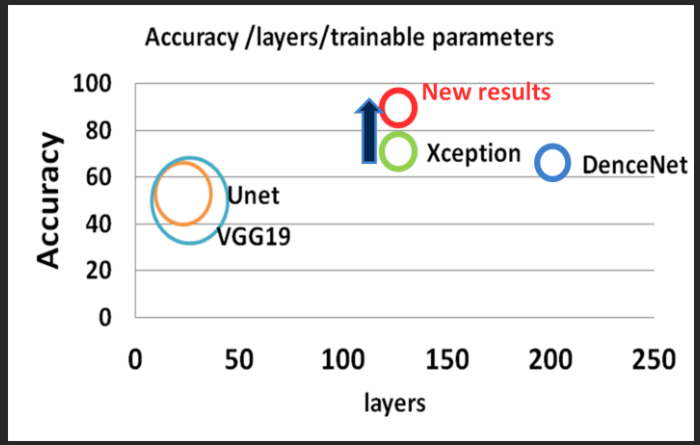


Correctly classified:

- Fractures: 27816 of the 32000
- Normal: 27603 of the 32000

- Sensitivity: 86.93%
- Specificity: 86.62%
- Accuracy: 86.6%
- AUC: 0.9264





Texas Children's