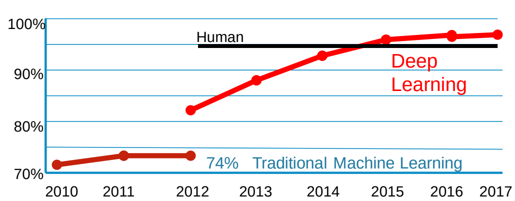
DEEP LEARNING WILL MAKE QUANTITATIVE IMAGING ROUTINE IN RADIOLOGY!

Bradley J Erickson, MD PhD
Mayo Clinic

DISCLOSURES

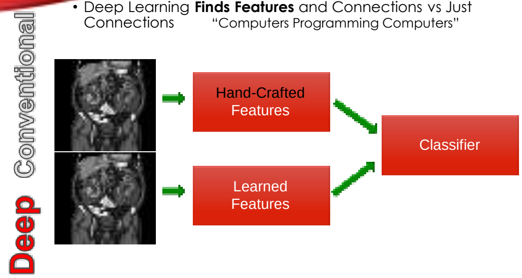
- Grant Funding: NIDDK DK-90728
- Commercial: none relevant to this work. Board positions in FlowSIGMA, VoiceIT, OneMedNet

DEEP LEARNING: WHY THE HYPE?



BENEFIT OF DL VS CONVENTIONAL ML

- Deep Learning **Finds Features** and Connections vs Just Connections "Computers Programming Computers"



MAN VS MACHINE

- **Too much hype** on computers replacing us
  - There is more variability in our job description than non-radiologists recognize (especially computer scientists).
  - Computers *will* start to 'read' screening/high volume exams and do mundane tasks (measurements) and create prelim reports

MAN VS MACHINE
MAN VS MAN + MACHINE

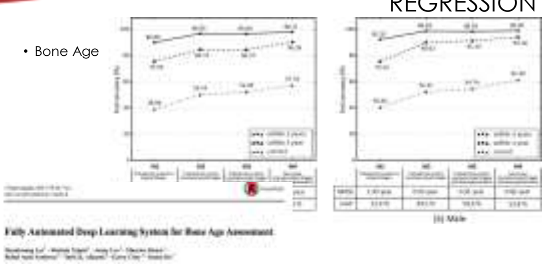
- **Too much hype** on computers replacing us
  - But they will do the routine/patterned tasks very well
- There is much more information in images that computers **can** extract
  - Quantitative values that are hard to get today
  - Unseen information that is lost

BASIC TASKS ML/DL CAN DO

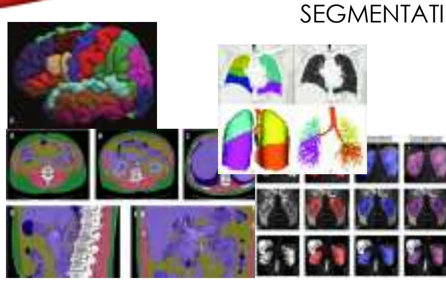
1. Regression
2. Segmentation
3. Classification
4. Detection (?Anomaly Detection)

REGRESSION

- Bone Age



SEGMENTATION



CURRENT RESULTS: BRAIN

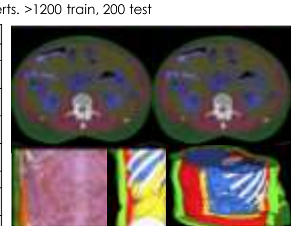
- Fully automatic segmentation of Brain
  - 127 parts of the brain semi-automatically labeled with human expert verification
  - 2000 train, 350 test
- Dice scores:
  - Mean: 0.954
  - Range: 0.852 – 0.989

MRI Truth Our Segmentation

CURRENT RESULTS: BODY COMPOSITION

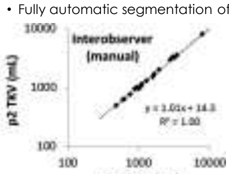
- CT slice at L3 traced by 2 human experts. >1200 train, 200 test

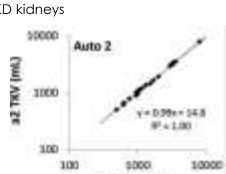
		Dice	Jaccard	TPF	FPF
Deep-learning vs gold-standard					
	SAT	0.98 (0.01)	0.96 (0.02)	0.98 (0.01)	0.02 (0.01)
	Muscle	0.95 (0.02)	0.91 (0.03)	0.95 (0.02)	0.05 (0.02)
	Viscera	0.99 (0.01)	0.98 (0.01)	0.99 (0.01)	0.01 (0.01)
	Bone	0.95 (0.02)	0.91 (0.04)	0.99 (0.01)	0.10 (0.05)
Manual inter-rater variability					
	SAT	0.95 (0.02)	0.91 (0.03)	0.94 (0.02)	0.03 (0.02)
	Muscle	0.93	0.87	0.93 (0.02)	0.07



CURRENT RESULTS: PKD

- Fully automatic segmentation of PKD kidneys



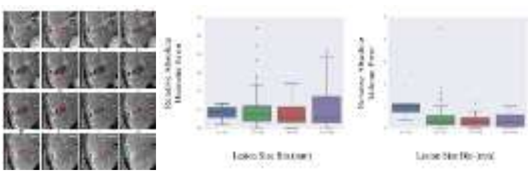


*Kline, J Digit Im, 2017

CURRENT RESULTS: CT ABDOMEN

- 100 CT abdomens for range of diseases
  - 20 were hand labeled by 5
  - 80 hand labeled by 1 (train)
  - STAPLE used to create 'truth' segmentation
- Dice scores for major organs:
  - Large globular organs like liver: 0.98
  - Thin plate-like organs like adrenal: 0.7
  - Remarkable resiliency to imaging technique
    - Recon kernel
    - Slice thickness
    - IV contrast
    - Oral contrast
  - Pathology had little impact (liver mets, massive ascites)

LIVER LESIONS

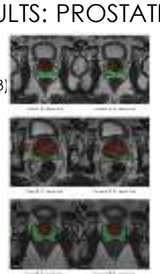


Sall, ISMRM, 2018

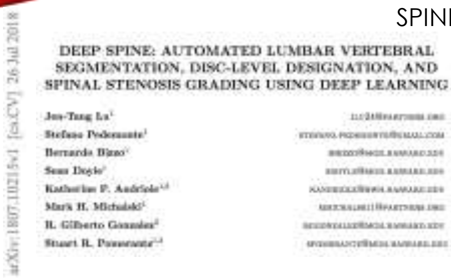
CURRENT RESULTS: PROSTATE

- 149 cases (125/24)
- Mean whole-gland Dice: 0.92 (95% CI 0.91-0.93)
- Mean zonal Dice: 0.88 (95% CI 0.87-0.89)

Bratt, ISMRM 2018



SPINE

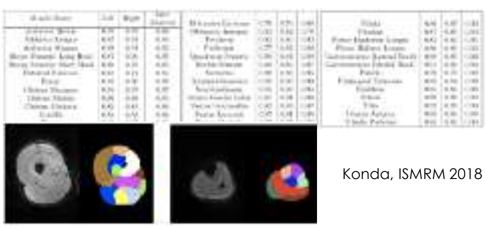


arXiv:1807.10215v1 [cs.CV] 26 Jul 2018

DEEP SPINE: AUTOMATED LUMBAR VERTEBRAL SEGMENTATION, DISC-LEVEL DESIGNATION, AND SPINAL STENOSIS GRADING USING DEEP LEARNING

Jen-Tang Lu[1]
Stefano Pedemonte[1]
Bernardo Bizzo[1]
Sean Doyle[1]
Katherine P. Andriole[1,2]
Mark H. Michalski[1]
R. Gilberto Gonzalez[2]
Stuart R. Pomerantz[1,2]

SEGMENTATION OF LEG MUSCLES!



Konda, ISMRM 2018

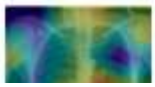
MEDICAL IMAGES WILL BE SEGMENTED

- Within 3 years, high volume CT and MR exams will be automatically segmented by deep learning
- Likely will be implemented on the scanner (5 years)
- Challenge: Standards (?IHE Profile) for conveying segmentations

CLASSIFICATION

- Is this disease X or Y? (Traditional CAD)

CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning



CLASSIFICATION

- Is this progression or response? (Emerging CAD)

SCIENTIFIC REPORTS

OPEN **Bladder Cancer Treatment Response Assessment in CT using Radiomics with Deep-Learning**

CLASSIFICATION

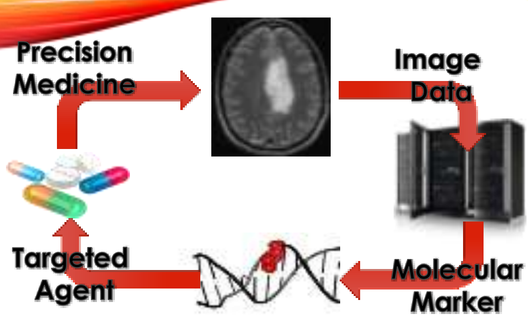
- Does the patient need treatment? (Emerging CAD)

kidney INTERNATIONAL

Image texture features predict renal function decline in patients with autosomal dominant polycystic kidney disease

CLASSIFICATION

- Is this disease X or Y? (Traditional CAD)
- Is this progression or response? (Emerging CAD)
- Molecular/Genomic Properties? (Super CAD)



PREDICTION OF 4 KEY MOLECULAR MARKERS FOR GLIOMA

N=498 subjects suing T2-weighted images from Mayo, UCSF, TCIA
398 for training, 100 for testing
50 layer ResNet, VGGNet, Inception

Marker	Sens	Spec	Accuracy
IDH1	0.95	0.95	0.95
1p19q Co-Del	0.91	0.85	0.87
ATRX	0.93	0.89	0.91
MGMT Methylation	0.95	0.95	0.95
Prediction of 'Triplet' (5 classes)	0.902	0.905	0.90

*Korfiatis, submitted

SCIENTIFIC REPORTS

OPEN

Predictive radiogenomics modeling of EGFR mutation status in lung cancer

Olivier Gevaert, Sebastian Echegaray, Amanda Khuong, Chuong D. Hoang, Joseph B. Shrager, Kirstin C. Jensen, Gerald J. Berry, H. Henry Guo, Charles Lau, Sylvia K. Plevritis, Daniel L. Rubin, Sandy Napel & Ann N. Leung

Results
Semantic image features show strong correlation with EGFR but not with KRAS mutation status.

RESEARCH ARTICLE

Radiogenomics of neuroblastomas: Relationships between imaging phenotypes, tumor genomic profile and survival

Hervé J. Brisse, Thomas Blanc, Gudrun Schleiermacher, Véronique Mosseri, Pascale Philippe-Chomette, Isabelle Janoueix-Lerosey, Gaelle Pierron, Eve Lapouble, Michel Peuchmaur, Paul Fréneaux, Louise Galmiche, Nathalie Algret, Matthieu Peycelon, Jean Michon, Olivier Delattre, Sabine Sarnacki

PLOS ONE | https://doi.org/10.1371/journal.pone.0185190 September 25, 2017

Conclusion

Imaging phenotypes of neuroblastomas, including tumor origin along the sympathetic system, correlate with tumor genomic profile and patient outcome.

AI AND MEDICAL IMAGING:

- Deep Learning applied to medical images can:
  - Do many mundane tasks like segmentation that may aid quantitative imaging
  - Predict important molecular markers with high accuracy
- Deep Learning applied to medical images may be a ***better*** tool for targeted therapies because it can:
  - Detect molecular properties, *and*
  - Determine how those are being expressed in that particular patient.

ISSUES

- Correlation does not prove causation
- Much of medicine accepts correlation, but must be validated
- Methods and data to validate correlations are difficult to obtain
- Methods to gain understanding from correlation are needed

THE REAL CHALLENGES TO DEEP LEARNING IN MEDICINE

- High Quality Annotated data sets
  - Large numbers may not be as important as high quality
  - Privacy concerns limit sharing
  - Annotation is a significant effort

THE REAL CHALLENGES TO DEEP LEARNING IN MEDICINE

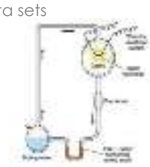
- High Quality Annotated data sets
- Diversity in data sets
  - There are known racial differences in disease appearance and response

THE REAL CHALLENGES TO DEEP LEARNING IN MEDICINE

- High Quality Annotated data sets
- Diversity in data sets
- Tools
  - Tools are very available.
  - Much published without adequate rigor

THE REAL CHALLENGES TO DEEP LEARNING IN MEDICINE

- High Quality Annotated data sets
- Diversity in data sets
- Tools
- Scientific Community
  - "There's no hypothesis"



CONCLUSIONS

- AI is and will continue to cause dramatic changes in the content of medical imaging and medical professionals
- We must remain engaged with technology development to assure it is implemented in a way that maximizes benefit to patients