Learning and Decision-Making from Big Data

Lei Xing, PhD, Jacob Haimson Professor

Departments of Radiation Oncology & Electrical Engineering
Stanford University

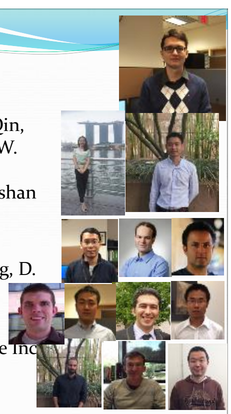
Stanford University
Department of Radiation Oncology
School of Medicine

Disclosures

- The Department of Radiation Oncology at Stanford University Hospital has a research agreement with Varian Medical Systems.
- Technology License Agreement with Varian. Dr. Xing has received speakers honoraria from Varian.
- Dr. Xing serves as advisory scientist in Zap Surgical, HuiyiHuiying Inc, and MoreHealth Inc.
- Research grants supports from NIH and Google

Acknowledgement

- Yixuan Yuan, Bulat Ibragimov, W. Qin, H. Liu, M. Korani, D. Li, K. Cheng, W. Zhao, C. Jenkins, S. Tzoumas, D. Vernekohl, S. Youselfi, , B. Ungan, Ishan Patil
- P. Dong, B. Han, A. Koong, D. Chang, D. Toesca, S. Soltys, R. Li
- Funding: NIH/NCI/NIBIB & Google Inc



Outline

- Introduction
- Analytics tools for "learning from data"
- Applications in medicine
  - Imaging
  - Treatment planning
  - Clinical studies
- Future outlooks and trends

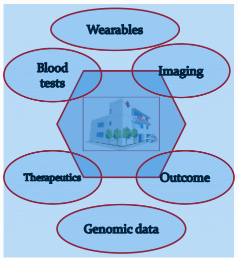
Big data is in our daily life and professional life

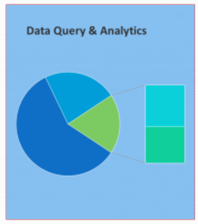
- Web sites: 40 billion indexed web pages
- Youtube: 100 hrs of videos are uploaded every minute
- WalMart: handels more than 1M transactions per hr.

Learning from data & examples

- A bank uses historical records of previous customers and yours to figure out a good formula for credit approval.
- A physician makes decision based on the patient's medical history and some symptoms.
- Predict how utility usage based on temperature and historical data.
- Interpolation and extrapolation scheme of experimental data and development of nomogram or empirical formula.

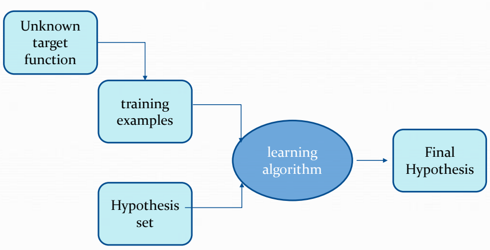
Big data in medicine, what and why?





How to make sense of big data?

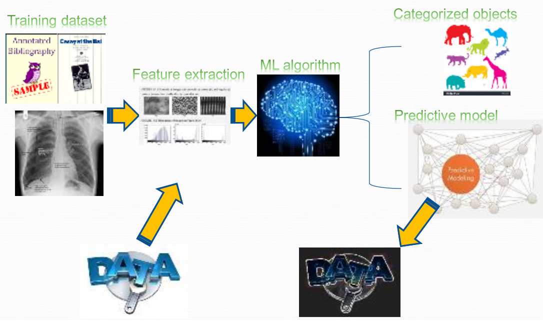
How do we learn from data?



Types of learning

- Supervised learning
- Unsupervised learning
- Reinforcement learning

Supervised learning



Issues related to learning from data

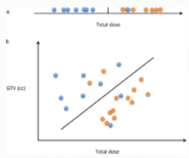
- What clinical problems can the approach solve?
- How big is big? Size of the data should not be the sole measure.
- It should be defined on a problem specific basis.
- Avoid dark data problem.

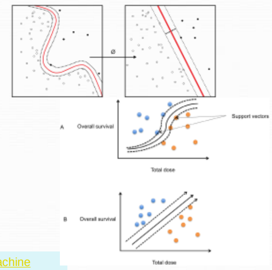
Computing & analytics tools are essential in dealing with big data

Machine leaning and deep learning

- Traditional statistical methods – logistic regression, Cox regression
- Machine learning methods –
  - decision trees (DT), a simple algorithm creates mutually exclusive classes by answering questions in a predefined orders;
  - Naïve Bayes (ND) classifiers, outputs probabilistic distributions among variables; k-nearest neighbors (k-NN);
  - Support vector machine (SVM), where a trained model will classify new data into categories;
  - Artificial neural net work (ANN), where models inspired by biological neural networks are used to approximate functions;
  - Deep learning, where multiple layers of neurons are used and is able to perform supervised and unsupervised learning.

Linear regression and support vector machine

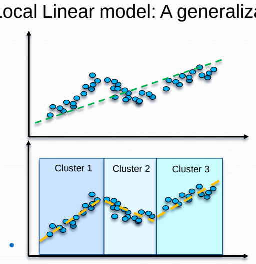




https://en.wikipedia.org/wiki/Support_vector_machine

P. Flach, Machine Learning, Cambridge Univ Press, 2012
J. Bibault et al, Big data and machine learning in radiation oncology: State of the art and future prospects, *Cancer Letters,* in press, 2016.

Local Linear model: A generalization to linear regression

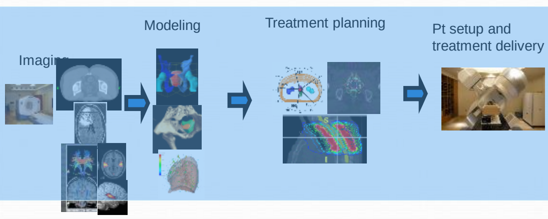


- Local linear model:
  - Step 1: Cluster data into subgroups
  - Step 2: Fit a linear model for each subgroup
- Dates back to 1970s (e.g., Diday, 1974)
- Local linear model is more general than the linear model (Hastie and Tibshirani, 1990)

- **Recent successful applications to:**
  - **Risk factors in heart disease (Jin and He, 2016)**
  - **Short-term load prediction in power systems (Dudek, 2016)**

Stanford MEDICINE | Radiation Oncology Medical Physics
Stanford | ENGINEERING Mechanical Engineering

Radiation Therapy Workflow



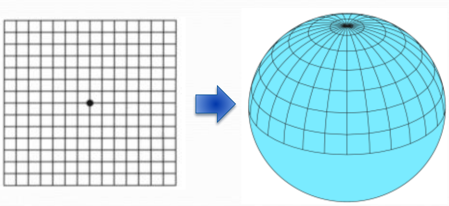
- Automation
- Artificial intelligence
  - Data, imaging, image guidance & integration

Types of learning

- Supervised learning
- Unsupervised learning
- Reinforcement learning

(input, some output, grade for this output)

Trajectory Optimization for VMAT



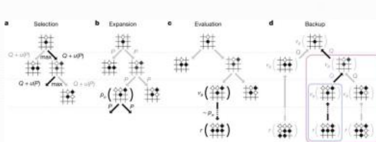
Decision: A trajectory of stone positions

Decision: A trajectory of Gantry angle, and Couch angle

Stanford MEDICINE | Radiation Oncology Medical Physics

Stanford | ENGINEERING Mechanical Engineering

Monte Carlo Tree Search (MCTS) Approach



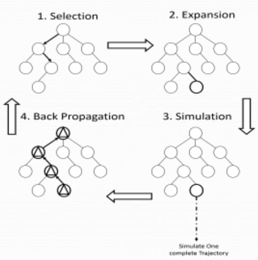
(Silver et al., 2016. Nature)

One of the essential components of AlphaGo is ***Monte Carlo Tree Search***

Stanford MEDICINE | Radiation Oncology Medical Physics

Stanford | ENGINEERING Mechanical Engineering

Monte Carlo Tree Search (MCTS) Approach

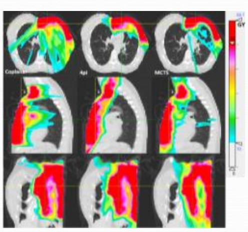


Monte Carlo tree search (Abramson, 1987): a heuristic search algorithm that expands a search tree based on random sampling of the search space.

Stanford MEDICINE | Radiation Oncology Medical Physics

Stanford | ENGINEERING Mechanical Engineering

Chest Wall Case Comparison: Coplanar, 4pi and MCTS



Compared with coplanar:

•*Heart mean dose:* reduced from 15 Gy to 7 Gy
Heart V30: reduced from 7.5% to 1.7%.
•*Left and right lung mean doses:* reduced from 16 Gy and 8 Gy to 11 Gy and 3 Gy, respectively

Compared with 4pi:

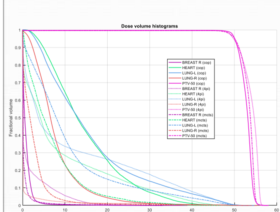
•MCTS spares spares more on ipsilateral lung, heart and contralateral breast.
•Contralateral lung is better spared in the 4pi plan.

Better quality than coplanar, competitive against 4pi

Stanford MEDICINE | Radiation Oncology Medical Physics

Stanford | ENGINEERING Mechanical Engineering

Chest Wall Case Comparison: Coplanar, 4pi and MCTS



Compared with coplanar:

•*Heart mean dose:* reduced from 15 Gy to 7 Gy
Heart V30: reduced from 7.5% to 1.7%.
•*Left and right lung mean doses:* reduced from 16 Gy and 8 Gy to 11 Gy and 3 Gy, respectively

Compared with 4pi:

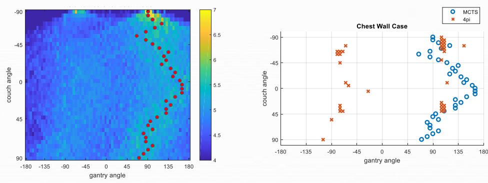
•MCTS spares spares more on ipsilateral lung, heart and contralateral breast.
•Contralateral lung is better spared in the 4pi plan.

Competitive against coplanar and 4pi

Stanford MEDICINE | Radiation Oncology Medical Physics

Stanford | ENGINEERING Mechanical Engineering

Trajectory optimized through MCTS (circle) and 4pi (cross)



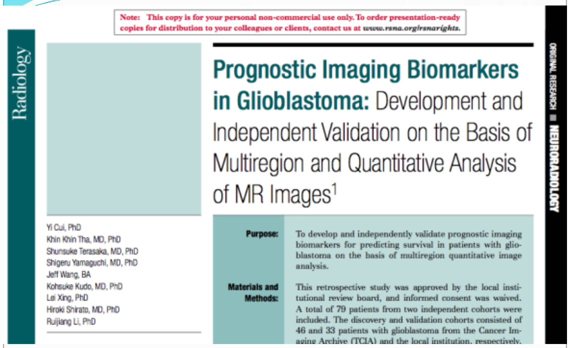
Conclusion

- Developed an AI-based trajectory design strategy for non-coplanar rotational arc radiation therapy
  - Flexibility: easily extendable to most *planning modules, optimization criteria*, or *computing architecture*
- A regularization scheme that encourages uniform fluence map
- Application of the proposed technique to two clinical cases suggests that it is capable of generating superior treatment plans and valuable for precision radiation therapy.

Radiomics for personalized medicine

- Biomarker: a measurable indicator of some biological state or condition
- Biomarker is a *key* element of personalized medicine.
  - Prognostic biomarkers: likelihood of disease progression – aggressive vs. indolent
  - Predictive biomarkers: sensitivity to therapy (drugs, radiation)
  - Early response biomarkers: spare patients ineffective treatment; speed up clinical trails.

Courtesy of Y. Cui



Radiology

ORIGINAL RESEARCH ■ NEURORADIOLOGY

Prognostic Imaging Biomarkers in Glioblastoma: Development and Independent Validation on the Basis of Multiregion and Quantitative Analysis of MR Images[1]

Yi Cui, PhD
Khin Khin Tha, MD, PhD
Shunsuke Terasaka, MD, PhD
Shigeru Yamaguchi, MD, PhD
Jeff Wang, BA
Kohsuke Kudo, MD, PhD
Lei Xing, PhD
Hiroki Shirato, MD, PhD
Ruijiang Li, PhD

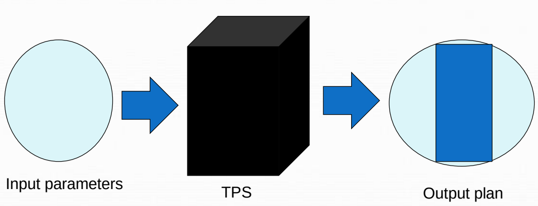
Purpose: To develop and independently validate prognostic imaging biomarkers for predicting survival in patients with glioblastoma on the basis of multiregion quantitative image analysis.

Materials and Methods: This retrospective study was approved by the local institutional review board, and informed consent was waived. A total of 79 patients from two independent cohorts were included. The discovery and validation cohorts consisted of 46 and 33 patients with glioblastoma from the Cancer Imaging Archive (TCIA) and the local institution, respectively.

Applications of big data in radiation oncology

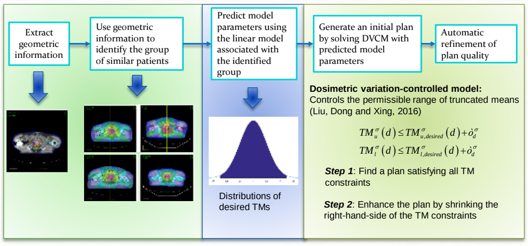
- Clinical studies.
- Radiomics (images are data!).
- Autopiloted and/or knowledge-based treatment planning.
- Machine learning/Deep learning for image registration and segmentation.

Inverse Treatment Planning



Problem: manual process with multiple trial and errors

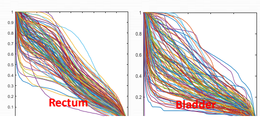
Workflow for a new patient

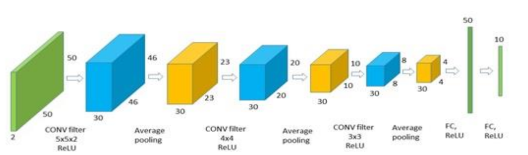


Stanford MEDICINE | Radiation Oncology Medical Physics
Stanford | ENGINEERING Mechanical Engineering

Deep learning for treatment planning

- prostate patient plans
  - Training dataset
  - 3 input channels (PTV, Blad., Rec.)
  - 8-layer convolutional network

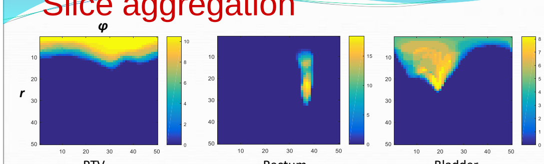




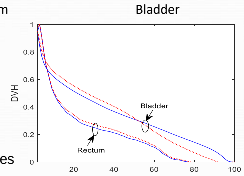
Moteza Korani

29

Slice aggregation

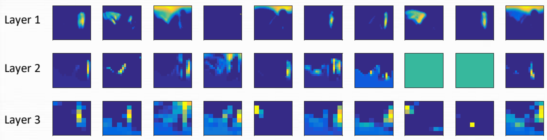


- Test performance
  - Normalized RMSE = 0.1 (Rectum)
  - Normalized RMSE = 0.2 (Bladder)
- *r*- and *φ*-aggregations may help!!!
  - Combine features from three branches

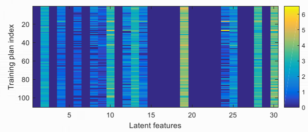


30

Feature maps



Only 16 active features!



31

Methods toward automated planning

- RapidPlan/Principle components
- Learning and deep-learning algorithm
- Multiobjective (RayStation)
- Automatic planning (Pinnacle)

Autopiloted treatment planning guided by big-data

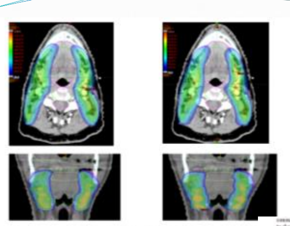
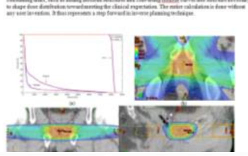


Fig. 8. Side-by-side comparison of the isodose distributions of autopiloted (right) and clinical (le



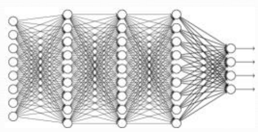
Applications of big data in radiation oncology

- Clinical studies.
- Radiomics (images are data!).
- Autopiloted and/or knowledge-based treatment planning.
- Machine learning/Deep learning for image registration and segmentation.

Deep learning

"A branch of machine learning based on a set of algorithms that attempt to model high-level abstractions in data by using multiple processing layers, with complex structures or otherwise, composed of multiple non-linear transformations."

Wikipedia

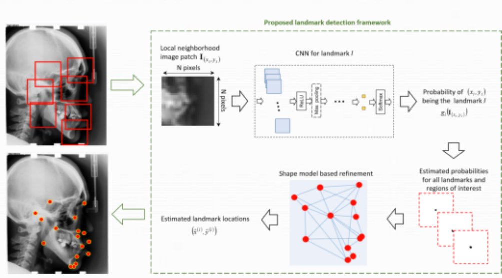


Major aspects of deep learning:

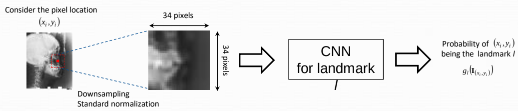
- Cascades multiple layers of processing units
- Units can comprise a broad family of linear/nonlinear functions for feature extraction and transformation
- Layers form a hierarchy from low- to high-level features
- Based on distributed representations assuming the observed data are generated by the interactions of factors

S Arik, B Ibragimov & L Xing

Landmark detection in cephalometric analysis



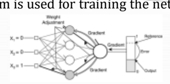
Overall framework for CNN- based landmark detection



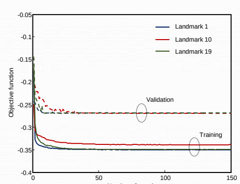
- Local spatial information of a pixel is used determine whether that pixel is one of the 19 landmarks.
- 19 different CNNs are trained, corresponding to 19 landmarks.
- Each CNN takes 34x34 patches as inputs and returns a scalar in [0,1], as a measure of the probability that pixel being the corresponding landmark.

Training the CNN

- Backpropagation algorithm is used for training the network.



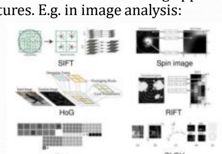
- Weights are initialized with small random variables and biases are initialized as zeros.



- The learning rate is initially chosen as 0.01 and exponentially reduced at each epoch in order to learn finer details.

Traditional machine learning vs. deep learning

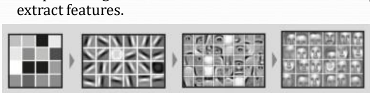
- Traditional machine learning approaches use hand-crafted relevant features. E.g. in image analysis:



Xie Y, Chao M, Xing L, Tissue feature-based and segmented deformable image registration for improved modeling of shear movement of lungs, Int J Radiat Oncol Biol Phys.74, 1256-65, 2009

Ref. *Deep learning methods for vision*

  - Non-flexible representation
  - Expert knowledge
  - Time-consuming hand-tuning
  - Problem specific
- Deep learning uses trainable feature extractors and optimizes the way to extract features.



Ref. *Convolutional deep belief networks, Honglak et al.*

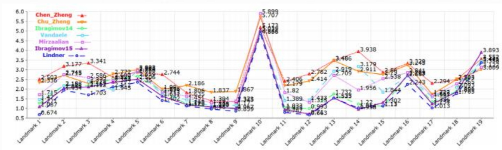
Landmark detection in cephalometric analysis

Technical (major) reasons for the choice of cephalometric analysis problem:

1)Availability of a sufficiently large image set (with ground-truth labels) in order to train a complex model

- 150 images for training and 250 images for testing

2)Availability of competitive benchmarks



Ref. Evaluation and Comparison of Anatomical Landmark Detection Methods for Cephalometric X-Ray Images: A Grand Challenge, C. Huang et al.

Clinical Application

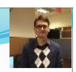
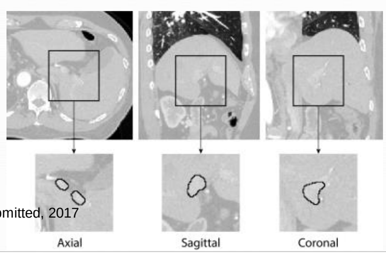


Image database

72 pre-treatment CT images:

- PV with contrast
- Stents
- Tumors are close to PV



B Ibragimov et al, submitted, 2017

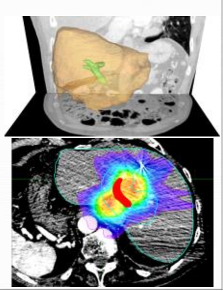
44

Central liver toxicity - B. Ibrimbrov, D. Toesca, A Koong, D. Chang, L XIng

Irradiation of hepatobiliary tract will likely result in central liver toxicity if the isotropic 15mm expansion of portal vein (PV) receives:

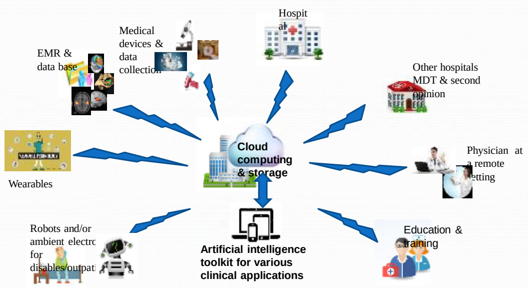
- $V_{BED10}30 > 45$ cc
- $V_{BED10}40 > 37$ cc

Can we predict such toxicity without manually annotating portal vein?



Stanford MEDICINE | Radiation Oncology Medical Physics

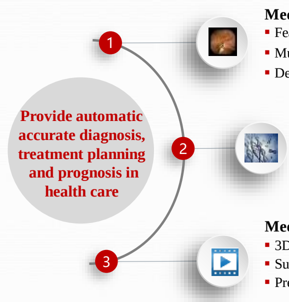
Telemedicine framework aided by AI



Summary

- What it takes for us to benefit from big data?
  - large database + analytics tools
  - Telemedicine infrastructure for big data-based medicine
- Machine learning and deep learning tools.

Future Work



Medical Images
- Feature extraction with domain knowledge
- Multi-modalities data analysis
- Deep learning application and modification

Biomedical Informatics
- High-level semantic feature extraction
- Feature fusion with image features
- RNN application

Medical Videos
- 3D reconstruction
- Surgical navigation
- Precision localization / Medical SLAM

Stanford University