Automation and Artificial Intelligence for Precision Radiation Therapy QA

Bin Han Ph.D.
July 31, 2018
AAPM Annual Meeting

Stanford University
Department of Radiation Oncology
School of Medicine

Disclosure

- I have no conflicts of interest to disclose.

Acknowledgement

Yong Yang, Ph.D.
Cesare H Jenkins, Ph.D.
Wei Zhao, Ph.D.
Yixuan Yuan, Ph.D.
Xiren Zhou
Lei Xing, Ph.D.
Karl Bush, Ph.D.
Nataliya Kovalchuk, Ph.D.
Amy Yu, Ph.D.

Stanford University

Outline

- Precision Radiation Therapy (PRT)
- Automation QA for PRT
- Artificial Intelligence (AI)
- AI applications in PRT QA

Stanford University

Precision Radiation Therapy

- Over decades of precision radiation therapy

CRT -> IMRT -> IGRT -> IMPT …….

- Future direction for precision radiation therapy:

More Imaging guidance modalities: MR PET and etc.
More adaptive radiation therapy via auto-replanning and checking.
More genomic and other prognosis take into account.

Challenges and Solutions

Challenges:

- Modern treatment machine: more components to QA
- Patient QA: treatment more complex & more adaptive plans
- Increasing physics chart checking and weekly chart QA.

Possible Solutions:

- Hire more physicists
- Automate the QA process
- Smart → AI

TG-142: A comprehensive Linac QA Guideline

- Dosimetry
- Mechanical
- Safety
- MLC
- Imaging: kV, MV, CBCT
- Respiratory gating
- Special procedures: IMRT/VMAT, SRS/SBRT, TBI,…
- Mordern Linac: 6D Couch, FFF beams and etc

Frequency:

- Daily, Monthly, Annually

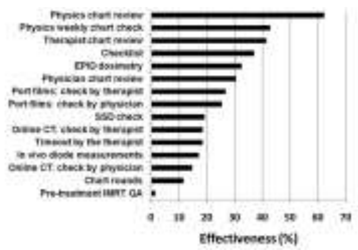
TG-218: IMRT Measurement-based Verification QA

Recommendation updates on:

- Different delivery methods
- Data interpretation
- Dose normalization
- Choice of tolerance limits for γ analysis
- Robustness analysis

QA Checks for Each Treatment

- QA checks are time consuming and prone to human errors.
- Some errors found after deliveries and may harm patients



Effectiveness: # of incidents that each QC check could detect/total # of incident reports.

Ford et al, Red Journal, 2012

Automated QA

- QA for a modern Linac for precision radiation therapy has been extremely extended with new components/functions added
- QA has become a complicated and very time consuming task

Table 1. Time (hours) spent undertaking linear accelerator QC testing

Time category	Minimum value	First quartile	Mean	Third quartile	Maximum value
Total machine time (hours per linear accelerator per month)	3.8	10.0	15.0	20.0	31.8
Total time including offline analysis (hours per linear accelerator per month)	5.8	13.7	[illegible]	30.2	56.8
Total time for patient-specific IMRT QC per patient	0.8	1.8	1.5	2.1	10.8

Automated QA: more efficient, stable and accurate

Palmer A et al, *Br. J. Radiol.* 2012(**85)** e1067–73

An Ideal Automatic QA Process for PRT

Automatic Delivery & Data Acquisition → Automatic Data Processing → Automatic Analysis & Reporting

Hardware + Software

- One button QA
- Self-calibration
- Phantom pose invariant
- Reduce/Remove operator dependence
- Analyze results and generate QA report

Automatic QA at Stanford

Direct visualization of Radiation

When radiation irradiates a radio-luminescent sheet fabricated from a mixture of GOS:Tb and PDMS, the irradiated area become visible.

Is this possible to use this to improve our QA processes?


Dr. Yang. Dr Jenkins

Courtesy of Cesare H Jenkins

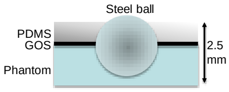
Jenkins C H *et al* 2015 *Med. Phys.* **42** 5–13

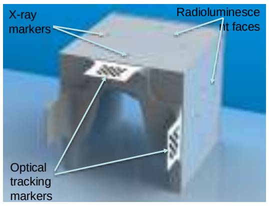
Automatic Mechanical QA

- Light Field/Radiation field coincidence
- Jaw position indicators
- Cross-hair centering
- Couch position indicators
- Laser localization

Phantom

- Structure fabricated on a MakerBot Z18 3D printer
- 2.38 mm stainless steel balls
- PDMS
- Gd_2O_2S:Tb





Jenkins C H et al Phys. Med. Biol. 61 (2016) L29

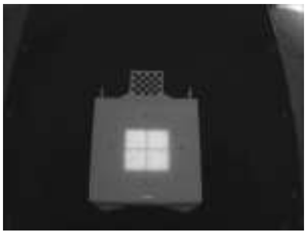
Camera

- Power over Ethernet (POE) machine vision camera
  - Single cable connection
  - 5mm f/2.5 S-mount lens
- 3D printed holder that connects to LINAC tray

Automatic Delivery/Operations

XML Script to implement:

- Turn on/off field light
- Set jaw positions
- Beam on
- Rotate gantry
- Turn on/off laser
- Treatment couch motions
- kV imaging
- Set MLC



Courtesy of Cesare H Jenkins

Image Processing

- Image identification and capture
- Transformation
- Analysis

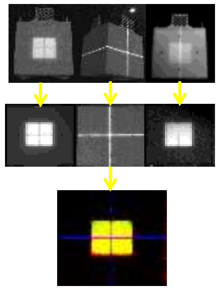
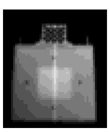


Image identification and capture

Key images were identified based on:

- Known delivery sequence
- Motion detection algorithm


Light Field

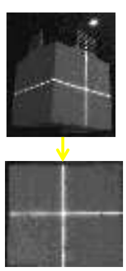

Radiation Field


Left Laser

Transformation

1. Transform the pixels corresponding to the phantom face into a calibrated image space
2. The transformation was determined as the linear transform that transforms the locations of the four fiducials to their aligned locations within the calibrated image space
3. The calibrated images were analyzed to identify the locations of salient features such as field edges, cross-hairs and lasers.

- *Self-calibration*
- *Correct for variations in setup*



Analysis

- Field Edges
  -Fit logistic function to find location of half value
- Crosshairs and lasers
  -Gaussian curve fitting
- kV and MV images
  -Image center is projected into the calibrated coordinate space

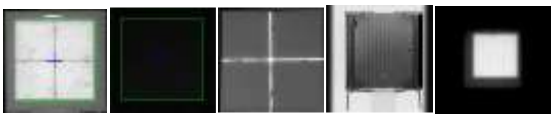
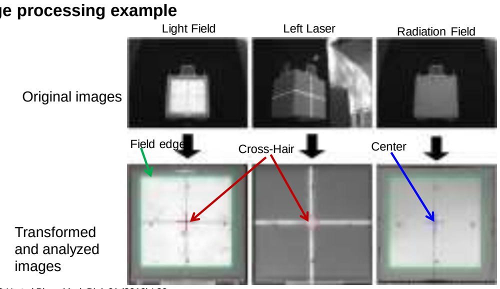


Image processing example



Jenkins C H et al Phys. Med. Biol. 61 (2016) L29

- Robust automated performance
- Accurate
  - ➢ *Be able to achieve 0.1mm~0.2mm accuracy, Better/Equivalent to current clinical practice*
- Repeatable
  - ➢ *Invariant to setup*
- More Efficient: ~10 min vs. manual 1~2 hours
  - ➢ *Set up: 7:00 min*
  - ➢ *Plan delivery: 1:21 min*
  - ➢ *Export DICOM: 1:00 min*
  - ➢ *Clean up: 2:00 min*

Artificial Intelligence and Machine Learning

- **Artificial intelligence (AI)** is intelligence demonstrated by machines. It perceives its environment and takes actions that maximize its chance of successfully achieving its goals. A machine mimics "**cognitive**" functions that humans associate with other human minds, such as "**learning**" and "**problem solving**".

- **Machine Learning** is an application of AI that provides systems the ability to **automatically learn** and **improve** from **experience** without being explicitly programmed.

Stanford University

Programing vs. Machine Learning

Traditional Programing

Data → Machine
Program → Machine
Machine → Model

Machine Learning

Data → Machine
Output → Machine
Machine → Model

Stanford University

Types of Machine Learning

Training Data

Supervised Learning: All Labeled → Model

Unsupervised Learning: All Unlabeled → Model

Reinforcement Learning: Reward or Penalty → Model

Stanford University

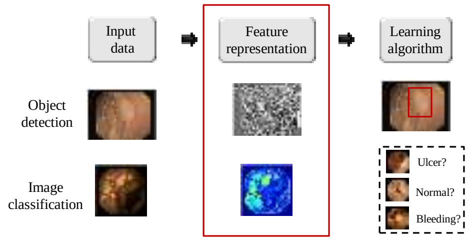
Machine Learning and Deep Learning

- Machine learning methods:
  - Linear Regression
  - Decision trees
  - Naïve Bayes classifiers
  - Support vector machine
  - Artificial neural net work
  - Deep learning

Stanford University

Machine Learning

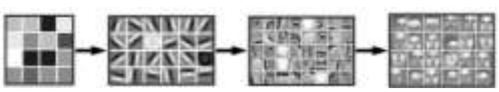
- Machine learning with hand-crafted features



Stanford University

Machine Learning

- Deep learning methods learn feature representations automatically
- Achieve good performance



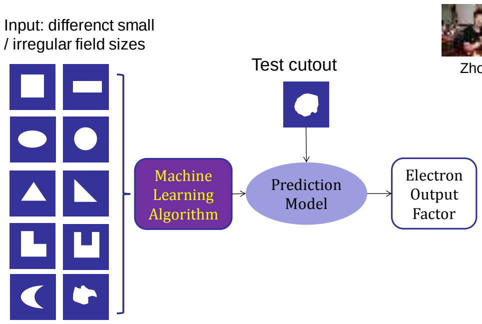
See (visual object recognition) Read (text understanding) Hear (speech recognition)

Stanford University

Machine Learning Applications in PRT @

- Electron small field output prediction
- Knowledge based chart checking

Electron Output Prediction




Zhou

Electron Output Prediction

- Scikit-learn package in python.
- Multivariate Linear regression with augmentation technique
- Total of 445 measurement data for training and testing
- The dose output factors for small and irregular electron treatment fields were accurately predict
- Mean relative absolute error 1.57%.
- R2 metric evaluation of the model is 0.994.

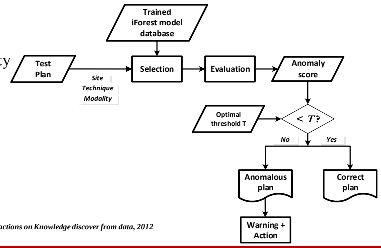
ML for Auto Treatment Plan Check


Dr Liu

❑ **Knowledge-based** error detection - learned from previous treatment parameters – anomaly detection

Isolation Forest:

- ➢ High detection sensitivity and specificity
- ➢ Linear computational complexity
- ➢ Low computer memory requirement



F. Liu, et al. Isolation Forest, ACM Transactions on Knowledge discover from data, 2012

Stanford University

Site, technique, modality dependent

Techniques	
SMLC	TANGENTS
2D	TBI
AP/PA	ELECTRON
	BOOST
3D	WEDGED PAIR
SBRT	CSA
DMLC	EN FACE

Major techniques
2D
IMRT
3D

Treatment modalities
MVX
Electrons

Treatment sites	
Brain	Lung
Pelvis	Breast
Head & Neck	Extremity
Thorax	Pelvic
Prostate	Chest wall
Abdomen	TBI
CSA	Lymph nodal
Skin	Enface

Stanford University

Data acquisition and pre-processing

❑ A total of 8335 patients with 11726 treatment plans since 2008 were acquired from R&V system (Dr. Shi Liu collected data from WUSTL).

❑ Parameters to be checked:

"Site", "Technique", "Modality", "Laterality", "SSD", "Fractions", "Fraction_dose", "Total dose", "MU", "MU/cGy", "Energy", "Beams", "Segments", "SegmentsPerBeam", "Use_MLC", "CPs", "CPsPerBeam", "MaxNCPs", "MinNCPs", "MaxMUcGyPerBeam", "MinMUcGyPerBeam", "PA", "PI", "PM", "PMU", and "PUAA".

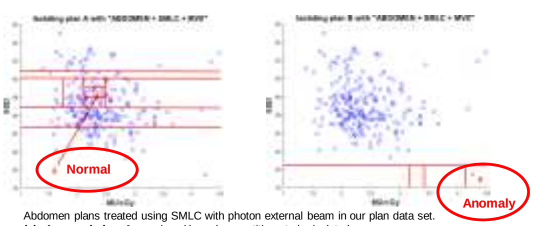
Stanford University

Plan parameters	Attributes	Data type	Attribute possible values
Treatment site	Site	Categorical	Brain, Breast, Prostate, …
Treatment technique	Technique	Categorical	2D, 3D, IMRT
Treatment modality	Modality	Categorical	MVX, Electron
SSD	SSD	Discrete	89.45 cm, 94.2 cm, 105 cm, …
Treatment fractions	Fractions	Discrete	2, 15, 25, 28, etc
Fractional dose	Fraction_dose	Discrete	180 cGy, 200 cGy, 800 cGy, …
Total dose	Total_dose	Discrete	4500 cGy, 5000 cGy, 5500 cGy, …
MU	MU	Discrete	247, 261, 226, …
MU/cGy	MU/cGy	Discrete	1.165, 1.27, 1.589, …
Energy	Energy	Discrete	6 MV, 10 MV, 6MeV, 9 MeV, …
Total # of beams	Beams	Discrete	1, 2, 4, 6, …

Plan parameters	Attributes	Data type	Attribute possible values
Total # of segments	Segments	Discrete	34, 44, 55, …
# of segment per beam	SegmentsPerBeam	Discrete	1, 6, 7, …
Use MLC or not	Use_MLC	Discrete	"1" or "0"
Total # of CPs	NCPs	Discrete	84, 98, 110, …
Minimum MU/cGy per beam	MinMUcGyPerBeam	Discrete	0.174, 0.344, 0.434, …
Plan averaged beam area	PA	Discrete	40.783, 56.259, … (set 0 for non-IMRT plans arbitrarily)
Plan averaged beam irregularity	PI	Discrete	4.684, 3.758 … (set 0 for non-IMRT plans arbitrarily)
Plan averaged beam modulation	PM	Discrete	0.596, 0.289, … (set 0 for non-IMRT plans arbitrarily)
Plan nomalized MU	PMU	Discrete	430, 474, … (set 0 for non-IMRT plans arbitrarily)
Plan averaged union area of all apertures	PUAA	Discrete	87.312, 136.908, … (set 0 for non-IMRT plans arbitrarily)

Stanford University

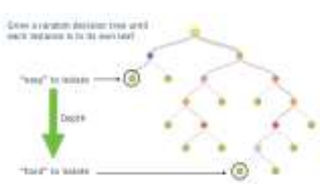
Detection method - iForest



Abdomen plans treated using SMLC with photon external beam in our plan data set.
(a) A normal plan, A, requires 11 random partitions to be isolated.
(b) An anomalous plan, B, requires only 4 random partitions to be isolated.

Stanford University

iForest – random tree structure



Training stage

- For each combination of plans with the same treatment site, technique and modality, an iForest model was trained by assembling a number of iTrees.
- Each iTree was built by recursively and randomly partitioning a sub-sample ($\phi = 256$) of the corresponding training data in each iForest model until all plans were isolated to leaf nodes.
- Inputs: training data set, number of trees t, and sub-sampling size ϕ, constituting the each iForest model, which is chosen by the user.
- Output: a dataset of trained iForest models, with each model consisting of plans with the same site, technique and modality

Swamping and masking

Figure. Illustration of swamping and masking effects introduced by large training data and reduced by sub-sampling, for a 2D (MU/cGy + SSD) distribution collected by all prostate plans treated using SMLC with photon external beam. Each blue dot represents a normal plan and each red dot represents an anomalous plan. (a) Original training data set including all the plans. (b) A sub-sample of the training data with fewer false positives.

Evaluation stage

- Purposely simulated various types of errors into our normal plan data set and compute the error detection rate using the trained models.

➢ **Median Absolute Deviation (MAD):**

$$MAD = bM_i(|x_j - M_j(x_j)|)$$

➢ **Error-level mechanism:**

50% and 100% error-levels simply represent that the introduced errors added are 50% and 100% of the original attribute-value, respectively.

➢ **Boxplot** $x_{min} < x < Q1 - 3IQR$, or $Q3 + 3IQR < x < x_{max}$

➢ **Manual addition: by a medical physicist**

Results – detection rates

Table 1. Averaged error detection results for different simulated error types

Error Types	Sensitivity (TPR)	Specificity (TNR)
MAD	99.53%	94.1%
100% Error-level	98.84%	91.8%
Boxplot	98.12%	90.2%
Manual addition	92.22%	87.6%

Results – error occurrences

Table 2. The most & least sensitive/specific parameters for different simulated error types

Error Types	Sensitivity		Specificity	
	Most	Least	Most	Least
MAD	PI (99.89%)	SSD (98.91%)	MaxN_CPs (96.55%)	SSD (90.21%)
100% Error-level	Total_dose (99.62%)	Energy (98.05%)	NCPs (95.91%)	SSD (86.27%)
Boxplot	NCPs (98.64%)	MU/cGy (97.92%)	Segments (94.37%)	MU (81.29%)
Manual addition	PA (95.56%)	Energy (90.05%)	PI (91.15%)	MU/cGy (79.76%)

Results – error occurrences

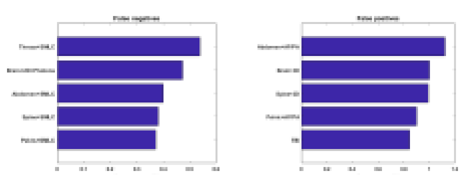
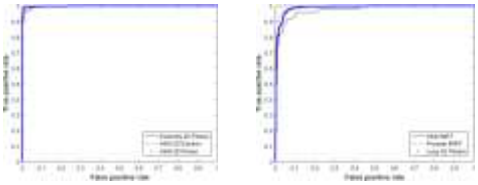


Figure. Histograms of obtained (a) false negative plans and (b) false positive plans with occurrences in percentages.

Results - ROC


(a) MAD errors simulation (b) 100% Error-level

High sensitivity and specificity to cover a wide range of treatment plan parameter errors.

Summary

- QA for precision radiation theapy has become a complicated and very time consuming task
- Automatic QA has the potential to provide QA procedures with high efficiency and less human error
- AI/machine learning is a rapid growing techniques and has the potential to optimize the future precision radiation therapy QA

Thank you very much!