Automation and Artificial Intelligence for Precision Radiation Therapy QA

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

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.

Outline

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

- Over decades of precision radiation therapy

\[\text{CRT} \rightarrow \text{IMRT} \rightarrow \text{IGRT} \rightarrow \text{IMPT} \ldots\]

- 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

- 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 \rightarrow 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,…
- Modern 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

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

Automated QA: more efficient, stable and accurate
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

Automatic Mechanical QA

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

Courtesy of Cesare H Jenkins
Phantom

- Structure fabricated on a MakerBot Z18 3D printer
- 2.38 mm stainless steel balls
- PDMS
- Gd$_2$O$_2$S:Tb

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

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

Original images

Transformed and analyzed images

- 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: 6:21 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.

---

Programming vs. Machine Learning

**Traditional Programming**

```
Data --> Machine --> Model

Program
```

**Machine Learning**

```
Data --> Machine --> Model

Output
```

---

Types of Machine Learning

**Training Data**

- **Supervised Learning**
  - All Labeled
  - Model

- **Unsupervised Learning**
  - All Unlabeled
  - Model

- **Reinforcement Learning**
  - Reward or Penalty
  - Model
Machine Learning and Deep Learning

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

Machine Learning

- Machine learning with hand-crafted features

Machine Learning

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

See (visual object recognition)  Read (text understanding)  Hear (speech recognition)
Electron small field output prediction

Knowledge based chart checking

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.
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

ML for Auto Treatment Plan Check

Site, technique, modality dependent

Techniques
- SMLC
- TANGENTS
- 3D
- TBI
- AP/PA
- ELECTRON
- BOOST
- SIBRT
- WEBDED-PAR
- DMLC
- EN FACE

Major techniques
- 2D
- MRT
- 3D

Treatment modality
- MAX
- Electrons

Treatment sites
- Brain
- Lung
- Pelvis
- Breast
- Head & NEck
- Extremity
- Thora
- Pelvic
- Prostate
- Chest wall
- Abdomen
- TBI
- CSA
- Lymph nodes
- Skin
- Enface

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:
### Plan parameters

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Data type</th>
<th>Possible values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site</td>
<td>Categorical</td>
<td>Brain, Breast, Prostate, ...</td>
</tr>
<tr>
<td>Treatment Site</td>
<td>Site</td>
<td>Categorical</td>
</tr>
<tr>
<td>Treatment Technique</td>
<td>Technique</td>
<td>2D, 3D, IMRT</td>
</tr>
<tr>
<td>Treatment Modality</td>
<td>Modality</td>
<td>MVX, Electron</td>
</tr>
<tr>
<td>SSD</td>
<td>Discrete</td>
<td>80.0 cm, 110.2 cm, 150 cm</td>
</tr>
<tr>
<td>Fractional dose</td>
<td>Fractional_dose</td>
<td>180 cGy, 200 cGy, 800 cGy, ...</td>
</tr>
<tr>
<td>Total dose</td>
<td>Discrete</td>
<td>4500 cGy, 5000 cGy, 5500 cGy, ...</td>
</tr>
<tr>
<td>MU</td>
<td>Discrete</td>
<td>247, 261, 226, ...</td>
</tr>
<tr>
<td>MU/cGy</td>
<td>Discrete</td>
<td>1.165, 1.27, 1.589, ...</td>
</tr>
<tr>
<td>Energy</td>
<td>Discrete</td>
<td>6 MV, 10 MV, 6 MeV, 9 MeV, ...</td>
</tr>
<tr>
<td>Total # of beams</td>
<td>Beams</td>
<td>1, 2, 4, 6, ...</td>
</tr>
<tr>
<td>Total # of segments</td>
<td>Segments</td>
<td>50, 60, 95, ...</td>
</tr>
<tr>
<td># of segment per beam</td>
<td>SegmentsPerBeam</td>
<td>1, 2, 3, 7, ...</td>
</tr>
<tr>
<td>Use MLC</td>
<td>Discrete</td>
<td>“1” or “0”</td>
</tr>
<tr>
<td>Total # of CPs</td>
<td>Segments</td>
<td>84, 98, 110, ...</td>
</tr>
<tr>
<td>Minimum MU/cGy per beam</td>
<td>MinMUcGyPerBeam</td>
<td>0.174, 0.344, 0.434, ...</td>
</tr>
<tr>
<td>Plan averaged beam area</td>
<td>PA</td>
<td>40.783, 56.259, ...</td>
</tr>
<tr>
<td>Plan averaged beam irregularity</td>
<td>PI</td>
<td>4.684, 3.758, ...</td>
</tr>
<tr>
<td>Plan averaged beam modulation</td>
<td>PM</td>
<td>0.596, 0.289, ...</td>
</tr>
<tr>
<td>Plan normalized MU</td>
<td>PMU</td>
<td>430, 474, ...</td>
</tr>
<tr>
<td>Plan averaged union area of all apertures</td>
<td>PUAA</td>
<td>87.312, 136.908, ...</td>
</tr>
</tbody>
</table>

### 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.

### 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 \( \phi \), 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 for all prostate plans treated using SMLC with planning-internal 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):**
    \[
    MAD = b M_f(|x_i - M_f(x_i)|)
    \]
  - **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**
    \[
    x_{min} < x < Q1 - 3IQR, \quad Q1 + 3IQR < x < x_{max}
    \]
  - **Manual addition:** by a medical physicist
### Results – detection rates

<table>
<thead>
<tr>
<th>Error Types</th>
<th>Sensitivity (TPR)</th>
<th>Specificity (TNR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAD</td>
<td>99.53%</td>
<td>94.1%</td>
</tr>
<tr>
<td>100% Error-level</td>
<td>98.84%</td>
<td>91.8%</td>
</tr>
<tr>
<td>Boxplot</td>
<td>98.12%</td>
<td>90.2%</td>
</tr>
<tr>
<td>Manual addition</td>
<td>92.22%</td>
<td>87.6%</td>
</tr>
</tbody>
</table>

### Results – error occurrences

<table>
<thead>
<tr>
<th>Error Types</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAD</td>
<td>99.53%</td>
<td>94.1%</td>
</tr>
<tr>
<td>100% Error-level</td>
<td>98.84%</td>
<td>91.8%</td>
</tr>
<tr>
<td>Boxplot</td>
<td>98.12%</td>
<td>90.2%</td>
</tr>
<tr>
<td>Manual addition</td>
<td>92.22%</td>
<td>87.6%</td>
</tr>
</tbody>
</table>

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

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

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

<table>
<thead>
<tr>
<th>Error Types</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAD</td>
<td>99.53%</td>
<td>94.1%</td>
</tr>
<tr>
<td>100% Error-level</td>
<td>98.84%</td>
<td>91.8%</td>
</tr>
<tr>
<td>Boxplot</td>
<td>98.12%</td>
<td>90.2%</td>
</tr>
<tr>
<td>Manual addition</td>
<td>92.22%</td>
<td>87.6%</td>
</tr>
</tbody>
</table>

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 therapy 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!