Associating 3D Radiation Dose with Treatment Failure in Prostate Cancer Radiotherapy Patients

Presented by Marco Marcello

MATLAB Conference 10/05/17
Acknowledgements

Supervisors – Professor Martin Ebert and Dr Pejman Rowshan Farzad

Trial Chairperson – Professor Jim Denham
Trial Contact – Mrs Allison Steigler
Breakdown

1) Introduction
   ⇒ Cancer and radiotherapy basics
   ⇒ Research problem and aims

2) Aim 1: Locating Regions where Dose is Correlated with Failure
   ⇒ Using MATLAB for 3D visualisation
   ⇒ Using MATLAB for locating significant regions

3) Aim 2: Classification Modelling
   ⇒ What is classification modelling?
   ⇒ Live demo of MATLAB Classification Learner App

4) Potential Impact of Research
   ⇒ How MATLAB helps impact radiotherapy through this project
What is cancer?

- Cancer results from genetic mutations that occur naturally or result from external influences.

- Particular mutations stimulate out-of-control cell growth of cells without normal functionality.

- Groups of these cells are what we call tumours, and can destroy healthy function in many crucial organs.
Cancer and Radiotherapy Basics

Is cancer prominent? How’s it treated?

- **Cancer** is the leading cause of death in Australia, with 1 in 2 men and 1 in 3 women diagnosed by the age of 85
- **Prostate cancer** is the most commonly diagnosed cancer in Australia, with 1 in 8 men diagnosed within their lifetime

Prostate cancer is typically treated with a combination of the following modalities:
Cancer and Radiotherapy Basics

What’s External Beam Radiotherapy (EBRT)?

**EBRT:** radiation (high energy photons) externally delivered to tumour via linear accelerator.

‘The Goal’: to maximise dose (energy deposition) to tumour and minimise dose to healthy tissues.
Prostate cancer external beam radiotherapy (EBRT) involves three major stages:
1) Produce a high-quality 3D image of the treatment region.
2) Produce a treatment plan, including a 3D radiation dose distribution.
3) Accurately deliver the planned radiation dose via linear accelerator.
Research Problem and Aims

Issues addressed? Project aims?

The project addresses a major unresolved issue in prostate EBRT:

⇒ The distribution of cancer around the prostate is difficult to locate due to potential microscopic disease spread.

⇒ Therefore, there could be regions that are under-dosed.

Project aims:

1) To find regions where dose variation is correlated with treatment failure.

2) To produce a classification model capable of predicting failure based on a subject dose distributions.

RADAR trial subjects

Aim 1: Locating Significant Regions

What’s a prostate EBRT 3D dose distribution?
Aim 1: Locating Significant Regions

3D Visualisation on MATLAB?
Aim 1: Locating Significant Regions

3D Visualisation on MATLAB?

Tools for NIfTI and ANALYZE image

version 1.27 (426 KB) by Jimmy Shen

Load, save, make, reslice, view (and edit) both NIfTI and ANALYZE data on any platform
Aim 1: Locating Significant Regions

Analysis: Dose-difference testing

RADAR trial subjects

1: Treatment failed
2: Treatment succeeded

\[ \bar{\mu}_1 \]
\[ \bar{\mu}_2 \]

\[ p < 0.01 \text{ threshold dose-difference map} \]

‘Locally normalised’ dose-difference map

Local anatomy

‘Multiple comparisons’ dose-difference test
Aim 1: Locating Significant Regions

Analysis: How did MATLAB help?

The entire analysis was performed on MATLAB

⇒ All dose distributions are 3D matrices
⇒ MATLAB’s matrix manipulation was very helpful
⇒ Using the ‘vec2mat’ function from the Communications System Toolbox

![Diagram showing matrix operations, statistical operations, and vec2mat function]
Aim 1: Locating Significant Regions

→ Analysis: How did MATLAB help?

The analysis involved a lot of data crunching!

⇒ Each dose distribution is about 2.3 Mb
⇒ The analysis includes over 680 distributions

Parallel Computing Toolbox allows running multiple core’s at once

⇒ Reduces processing time, very helpful!
⇒ Especially ‘for loops’, frequently used in code
Aim 2: Classification Modelling

What is ‘machine learning’ based classification?

A number of machine learning algorithms are used for classification:

- K-nearest neighbour
- Support vector machine
- Decision tree

Two major steps:

1) Use an algorithm to ‘train’ a model using a data set made up of classes (subgroups).
2) Classifying external data into a class of the training set.
Aim 2: Classification Modelling

What is ‘machine learning’ based classification?

Class 1

Class 2

Training set

Machine learning algorithm

Classification model

Test data

Class 1
Aim 2: Classification Modelling

What is ‘machine learning’ based classification?

How do we test our model without test data? Can use cross-validation...

Class 1

Class 2

Training set

Machine learning algorithm

Classification model

Test data

Class 1
Aim 2: Classification Modelling

What is ‘machine learning’ based classification?

Three basic statistics help us test the model’s performance:

Accuracy:
\[ Acc = \frac{\text{# correctly classified}}{\text{total number}} \]

E.g. 10 cats, 10 dogs
7 cats and 9 dogs correctly classified
\[ \Rightarrow Acc = \frac{7 + 9}{20} = 80\% \]

Sensitivity:
\[ Sen = \frac{\text{# correctly classified}}{\text{total number in class 1}} \]
\[ \Rightarrow Sen = \frac{9}{10} = 90\% \]

Specificity:
\[ Spe = \frac{\text{# correctly classified}}{\text{total number in class 2}} \]
\[ \Rightarrow Spe = \frac{7}{10} = 70\% \]
E.g. Classifying RADAR data according to number of beams:

Not 4 beams

4 beams

Class 1

Class 2

Training set

Machine learning algorithm

Classification model

Class 1

Not 4 beams
Aim 2: Classification Modelling

Benefits of the MATLAB Classification Learner App

Benefits:

- Running multiple machine learning algorithms simultaneously is extremely powerful
- Implement cross validation easily
- Activate PCA easily
- Generate performance statistics easily
Aim 2: Classification Modelling

Class 1: Treatment ‘failed’
(passed away)

Class 2: Treatment ‘succeeded’
(survived)

Goal: to build a model capable of predicting treatment failure in a patient based on their planned dose distribution.
Aim 2: Classification Modelling

Classification results:

- Classified the two groups:
  - 197 died within 6.5 years
  - 487 death free
- 1 in 8 sampled data
- 15-fold cross validation

Model 1.7

![ROC curve](image)

AUC = 0.67

Accuracy:

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Tree</td>
<td>68.4%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.2 Tree</td>
<td>70.8%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.3 Tree</td>
<td>71.9%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.4 Linear Discriminant</td>
<td>67.4%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.5 Quadratic Discriminant</td>
<td>67.4%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.6 Logistic Regression</td>
<td>48.4%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.7 SVM</td>
<td>72.4%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.8 Linear SVM</td>
<td>70.5%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.9 Cubic SVM</td>
<td>69.2%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.10 Fine Gaussian SVM</td>
<td>70.6%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.11 Medium Gaussian SVM</td>
<td>72.4%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.12 Coarse Gaussian SVM</td>
<td>71.1%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.13 KNN</td>
<td>65.8%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.14 KNN</td>
<td>71.8%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.15 KNN</td>
<td>71.9%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.16 KNN</td>
<td>71.1%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.17 KNN</td>
<td>60.7%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.18 KNN</td>
<td>70.5%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.19 Ensemble</td>
<td>70.0%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.20 Ensemble</td>
<td>72.1%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.21 Ensemble</td>
<td>60.5%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.22 Ensemble</td>
<td>63.9%</td>
<td>10752/10752 features</td>
</tr>
<tr>
<td>1.23 Ensemble</td>
<td>66.5%</td>
<td>10752/10752 features</td>
</tr>
</tbody>
</table>
Potential Impact of Research

How can these MATLAB based analyses impact radiotherapy?

Impact of these analyses:

1) Identifying anatomical regions prone to under-dosing is helpful because:
   ⊳ Valuable for clinicians identifying the optimal treatment region
   ⊳ It helps determine where and why the treatment is failing
   ⊳ Correcting this can directly improve patient outcomes!

2) A robust model capable of predicting treatment failure is helpful because:
   ⊳ It can help us determine particular geometric dose patterns associated with treatment failure
   ⊳ Predicting treatment failure before treatment can indicate need for replanning and improvement
Questions...