

Deep Learned and Practical Implementing Machine Learning Techniques in the Real World

Timor Kadir DPhil

Mirada Medical CTO
Optellum CTO
Visiting Fellow University of Oxford

7/31/2017

1

Disclosures

- Mirada Medical
 - Employee
 - Shareholder
- Optellum
 - Employee
 - Shareholder

Overview

- I – How do we bring new algorithms to the clinic?
- II – Practical challenges
 - Automatic OAR Segmentation (Mirada)
 - Lung Nodule Risk Stratification (Optellum)
 - Oxford SpineNet (Oxford University)
- III - Perspectives: potential impact of ML in Radiation Therapy.

7/31/2017

1

Timor Kadir

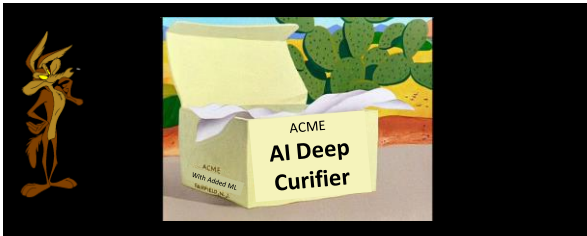
1992 - Electrical and electronic engineering at Surrey University
1995 - Motorola working on mobile video and image coding for 3G
1998 - DPhil and post-doc at Oxford Robotics Research Group
2004 - Medical image understanding @ Siemens Molecular Imaging
2009 - CSTO @ Mirada Medical / Visiting Fellow Oxford
2016 - Co-founder & CSTO @ Optellum

Logos: UNIVERSITY OF SURREY, MOTOROLA, SIEMENS, MIRADA medical, Optellum

Medical Devices: From bench to bedside



Medical Devices: From bench to bedside



From bench to bedside



From bench to bedside

Data Collection	Lung nodule risk stratification
Training	Spinal MRI analysis
Validation	Automatic contouring

Early Detection by Chest Computed Tomography Significantly Improves Lung Cancer Survival

• Lung cancer can initially present as a Pulmonary Nodule (PN).

April 2015: Nodules detected in two patients. Recommended CT follow-up to monitor growth.

Problem: In Current Clinical Practice, Pulmonary Nodules are Guilty and Followed Up Excessively Until Proven Innocent.

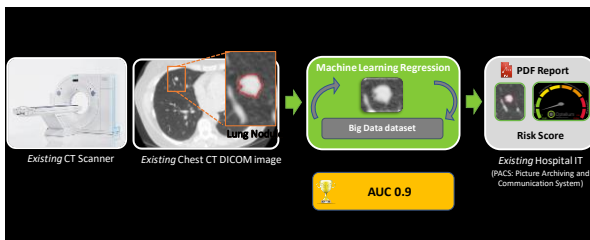
97% false positive rate
10M new patients - 10M new years of anxiety p.a.
20M new CT for low-ups p.a.
Extensive morbidity (radiation, biopsy, surgery)

1-Year Follow-up: Benign (L, majority of cases), Malignant (M, minority of cases)

7/31/2017

50

Nodule Risk Stratification



7/31/2017

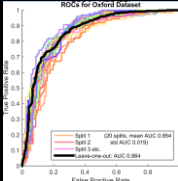
51

The ugly truth....

- Yes we won LungX.
- But the AUC was only ~0.65
- AUC on training data was >0.85
- Why the difference?
- Training Data

Look a little closer

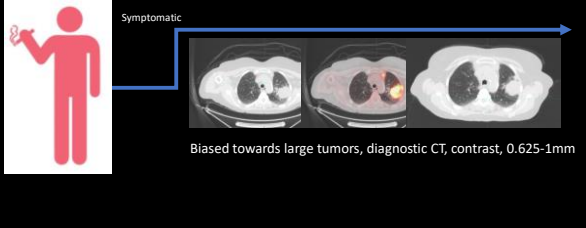
- LungX
 - 10 "calibration" datasets with known ground-truth
 - 100 test set with unknown ground-truth
- We manually created a dataset from Oxford and LIDC data trying to match the characteristics of the LungX datasets.
- Bias – in LIDC most malignant lesions are large and the benign ones are small.



The figure is a Receiver Operating Characteristic (ROC) curve titled "ROCs for Oxford Dataset". The y-axis is labeled "True Positive Rate" and the x-axis is labeled "False Positive Rate", both ranging from 0 to 1.0. The plot shows four curves representing different splits of the dataset. A legend in the bottom right corner provides the following data:

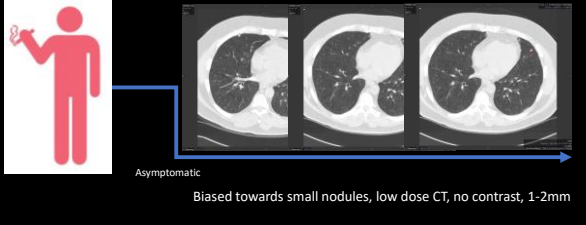
Split	GT splits	mean AUC
Split 1	10	0.894
Split 2	10	0.899
Split 3	10	0.894
Split 4	10	0.894

Dataset biases



The diagram illustrates a "Symptomatic" patient, represented by a red stick figure holding a hand to their chest. A blue arrow points from the figure to three axial CT scan slices of the chest. The scans show large, prominent lung tumors. Below the scans, the text reads: "Biased towards large tumors, diagnostic CT, contrast, 0.625-1mm".

Dataset biases



The diagram illustrates an "Asymptomatic" patient, represented by a red stick figure holding a hand to their chest. A blue arrow points from the figure to three axial CT scan slices of the chest. The scans show small, subtle lung nodules. Below the scans, the text reads: "Biased towards small nodules, low dose CT, no contrast, 1-2mm".

Timor's recipe #1 for ML success

- Acquire data that represents the typical distribution of disease, patient and imaging characteristics
- Test for built-in biases
- If you are going to use heterogeneous data for training, at least test with realistic data



From bench to bedside

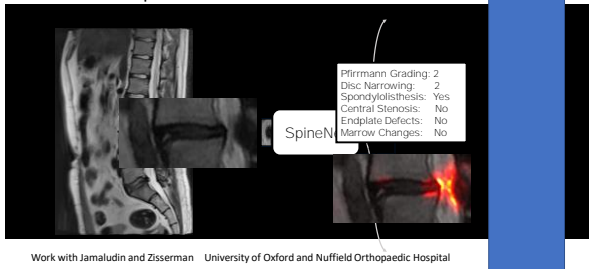
- Data Collection
- Training
- Validation

From Computer Vision

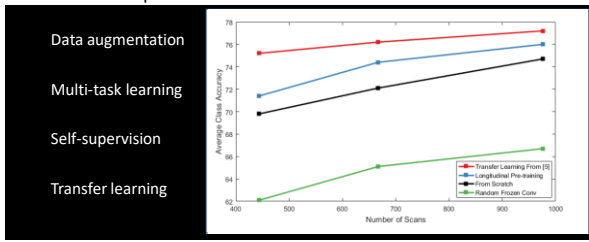
- ImageNet 2017 Object Localisation Challenge
 - 1000 objects: 100k/50k training and test images
 - Winner used Res-Net with > 150 layers
- Strategies for reducing training requirements
 - Data augmentation
 - Transfer learning
 - Multi-task learning
 - Self-supervision



Oxford SpineNet



Oxford SpineNet



Timor's recipe #2 for ML success

- Explore the utility of transfer learning, dataset augmentation, self-supervision and multi-task learning
- Other techniques are available: e.g. adversarial training
- Test robustly: if the distribution is not in the training data then it won't learn

From bench to bedside

Data Collection

Training

Validation

Auto-contouring Quality Assessment

Quantitative

Qualitative

Turing Test

Atlas-based

Learn the Details | Overview | Participants | Results | Summary

Challenge | **The AAPM Challenge** | **Medical Challenge**

PT4246 - 05/01/2017
 Participants perform segmentation on head CT. We will share and submit their results to compete for the three weeks to present their work at Challenge Competition at the AAPM 2017 Annual Meeting.
 Max submissions per day: 3
 Max submission size: 100M

Download CSV

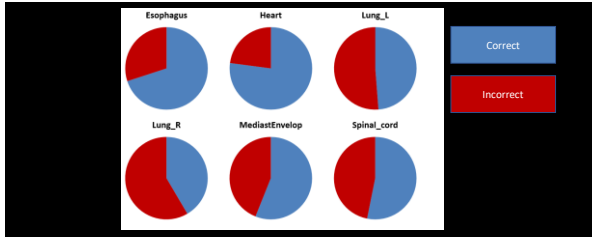
Rank	Team	Score	Target Accuracy	Team Name	Challenge #	Download Score
1	medrxiv	1	0.971817		10,262,010 (1)	View
2	panabdur	2	0.971817		10,262,010 (2)	View
3	mlabing	3	0.971817	uab-ill	1,010,571 (3)	View
4	bruno	4	0.971817		16,432,047 (4)	View
5	jam	5	0.971817		41,632,788 (5)	View
6	medrxiv	6	0.971817	Uls, Pathology	16,228,067 (6)	View
7	medrxiv	7	0.971817		10,262,010 (7)	View
8	medrxiv	8	0.971817		10,262,010 (8)	View
9	medrxiv	9	0.971817	Zhuo,UMC	11,272,010 (9)	View
10	medrxiv	10	0.971817		11,254,499 (10)	View

AAPM Auto-segmentation Grand Challenge Thursday 10am

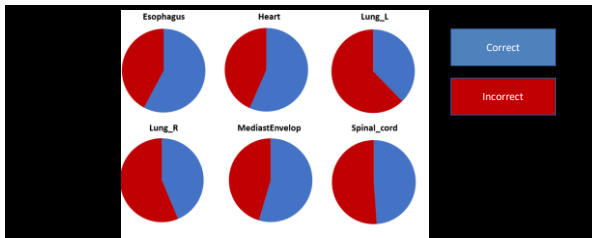
Turing Test

Turing test for contouring: www.autocontouring.com

Atlas-based contouring



DLC



Timor's recipe #3 for ML success

- Conventional methods for validation should be utilized for your proposed method
- Consider Turing test type approaches
- Don't rely on one validation approach

The image contains two smaller images. The top one shows a golden, stylized robot head with a human-like face. The bottom one shows a man in a blue shirt and apron cooking with two children, a boy and a girl, in a kitchen.

Perspectives

Machine Learning is in its infancy in Radiation Therapy

Searched on: CNN, Machine Learning, Deep Learning, Convolutional
71 abstracts

Topics include:

Segmentation, Classification, Treatment Planning, DIR, Response
assessment, Dose estimation, Image reconstruction.....

Perspectives

Data is critical but beware hidden biases

Lots of techniques now available to maximise power of datasets

Think broadly about validation protocols: Turing test
