Machine Learning to Improve Human Learning (or Understanding) from **Longitudinal Image Sets**

Harini Veeraraghavan
Memorial Sloan Kettering Cancer Center

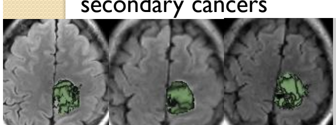
Collaborators

Joseph O. Deasy, PhD
Duc Fehr, PhD
Evis Sala, MD PhD
Alberto H. Vargas, MD
Hedvig Hricak, MD PhD
Alexandra Snyder-Charen, MD
Robert Young, MD
Ingo Mellinghoff, MD

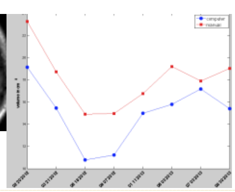
Elizabeth Sutton, MD
Elizabeth Morris, MD
Nancy Lee, MD
Neelam Tyagi, PhD
Margie Hunt, PhD
Jung-Hun Oh, PhD
Aditya Apte, PhD

Why Longitudinal Radiomics?

- Understanding how patients' disease progresses over time
- Assessing patients' response to treatment
- Identifying patients at risk of developing secondary cancers



02-20-2012 19.143 cm³ **09-07-2012 11.238 cm³** **08-30-2013 17.260 cm³**



Issues With Robust Longitudinal Radiomics Analysis

- Missing and inconsistent data
  - Not all data are created equal
  - Variability in the appearance, presence/absence of structures of interest with inter and intra-patients
- Highly unbalanced datasets
- Segmentation of structures of interest

Solutions

- Highly variable datasets
  - Novel representations of the data
  - Learning to deal with missing data
- Highly unbalanced data
  - Sample augmentation-based machine learning
- Segmentation
  - Machine-learning and semi-automatic longitudinal image segmentation methods

Using Segmentation For Radiomics

ORIGINAL RESEARCH

Breast Cancer Subtype Intertumor Heterogeneity: MRI-Based Features Predict Results of a Genomic Assay

Elizabeth J. Sutton, MD,[1*] Jung Hun Oh, PhD,[2] Brittany Z. Dashevsky, MD, PhD,[1,3] Harini Veeraraghavan, PhD,[2] Aditya P. Apte, PhD,[2] Sunitha B. Thakur, PhD,[1,2] Joseph O. Deasy, PhD,[2] and Elizabeth A. Morris, MD[1]

OPEN ACCESS Freely available online PLOS | ONE

Robust Radiomics Feature Quantification Using Semiautomatic Volumetric Segmentation

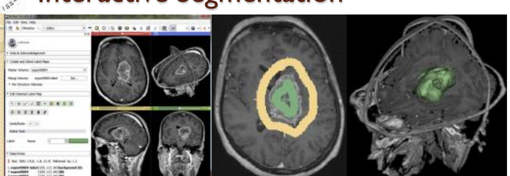
Chintan Parmar[1,2,3*], Emmanuel Rios Velazquez[1,2], Ralph Leijenaar[2], Mohammed Jermoumi[1,4], Sara Carvalho[2], Raymond H. Mak[1], Sushmita Mitra[3], B. Uma Shankar[3], Ron Kikinis[5], Benjamin Haibe-Kains[6,7], Philippe Lambin[1], Hugo J. W. L. Aerts[1,2,5*]

1 Department of Radiation Oncology, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, 2 Department of Radiation Oncology (MAASTRO), Maastricht University, Maastricht, The Netherlands, 3 Machine Intelligence Unit, Indian Statistical Institute, Kolkata, India, 4 University of Massachusetts, Lowell, Massachusetts, United States of America, 5 Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, 6 Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada, 7 Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

Issues with Manual Segmentation

- Manual segmentation is
  - Highly accurate (most of the times)
  - Time consuming and labor intensive,
  - Highly variable

Less Labor-Intensive Approach: Interactive Segmentation



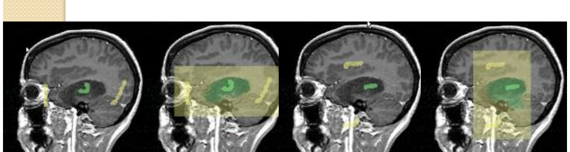
User Interface | Mark Target and Background scribbles | Generated Segmentation

Grow Cut Segmentation

J. Egger, T. Kapur, A. Fedorov, S. Pieper, J.V. Miller, H. Veeraraghavan, B. Friesleben, A.J. Golby, R. Kikinis, "GBM volumetry using 3DSlicer medical image computing platform", Sci Reports, 2013

What about Repeatability?

- Achieving **objective segmentation** in **real-time** with **repeatability** is difficult
  - Automatic methods are repeatable but less accurate
  - Semi-automatic (interactive) methods are highly accurate but not repeatable

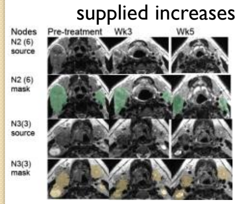


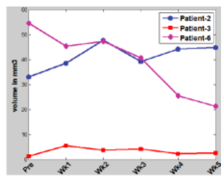
Case 1 | Case 2

9

Longitudinal Segmentation

- Requires user interaction at each time
  - Can be rather painful when a number of strokes need to added, corrections need to supplied increases





15 mins for each time point => 5 time points (75 mins)!!

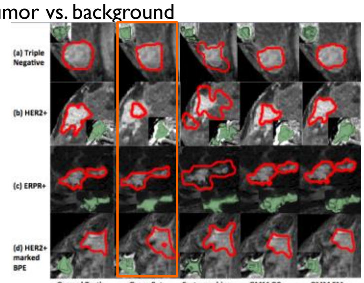
H.Veeraraghavan, N.Tyagi, M. Hunt, N. Riaz, S.McBride, N.Lee, J.O.Deasy, AAPM 2014

Solutions

- Combining machine learning to **reduce user interactions**
  - Algorithm learns model of target from user strokes and segments
  - Algorithm generates queries to improve segmentation online – active learning
- Fully automatic:
  - Combining machine learning with atlas

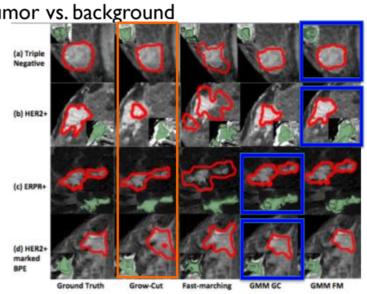
Combining Machine Learning With User Input

Gaussian mixture model (GMM)-based learning of tumor vs. background

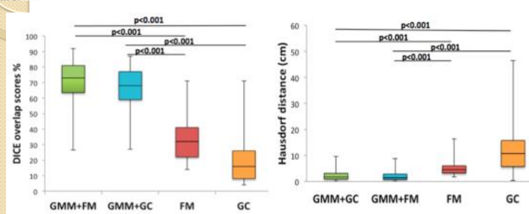


Combining Machine Learning With User Input

Gaussian mixture model (GMM)-based learning of tumor vs. background



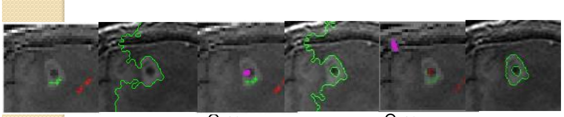
Adding Learning also Reduces Variability in Segmentation Accuracy



H.Veeraraghavan,E.J.Sutton, B.Dashevsky, E.Morris, J.O.Deasy. In preparation

Active Learning with Interactive Segmentation

- Active learning generates queries to improve the segmentation accuracy automatically
- Learning **guides user interactions** to achieve repeatable segmentation

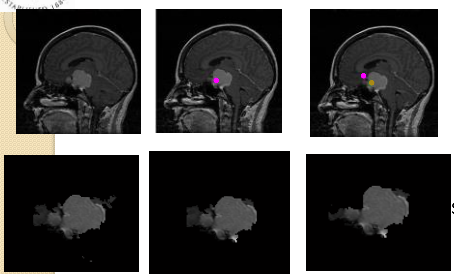


User marked | Query suggested | Query suggested

H.Veeraraghavan, J.V. Miller, "Active learning guided user interactions for consistent image segmentation with reduced user interactions", ISBI 2011

15

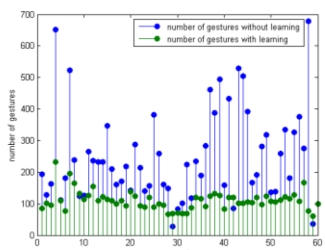
Algorithm Queries with Segmentations



Queries

Segmentations

Number of User Interactions Required with Learning is Significantly Lower than When Not Using Learning



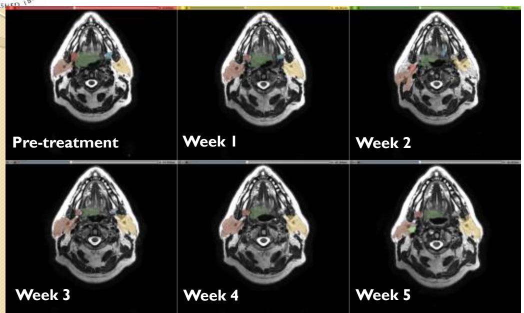
Increasing accuracy with each suggestion

Iterations	1	2	3	4	5
p-value <	0.006	0.001	0	0	0
p-value ≠	0.011	0.002	0.001	0	0
t-value	-2.71	-3.42	-3.77	-4.72	-5.13

Fully Automatic Longitudinal Segmentation

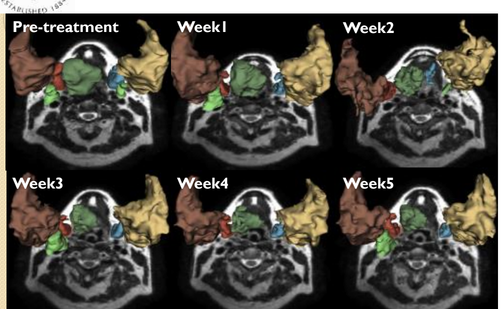
- Atlas or Patient specific segmentation
  - Involves an image registration to a patient or multi-atlas
  - Refine segmentation from the atlas
    - Machine learning-based classification (optionally) followed by volumetric segmentation

Longitudinal Segmentation (Pat 1)



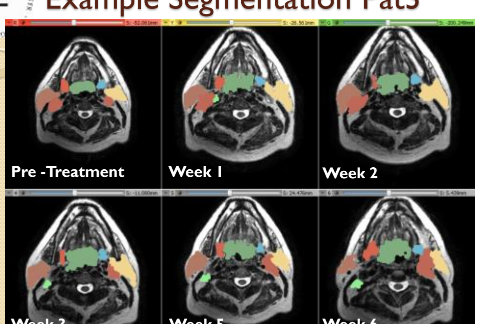
AutoSeg 2015

Longitudinal Segmentation (Pat 1)



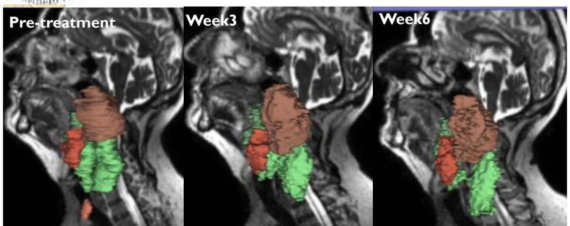
AutoSeg 2015

Example Segmentation Pat3



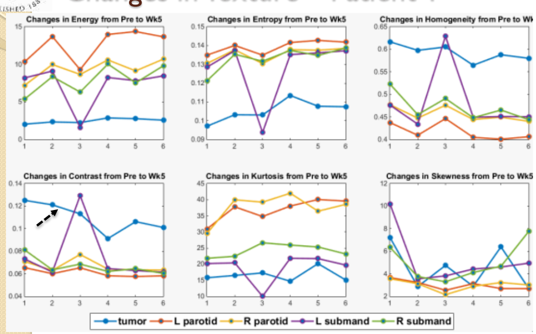
AutoSeg 2015

Example Segmentation Pat3

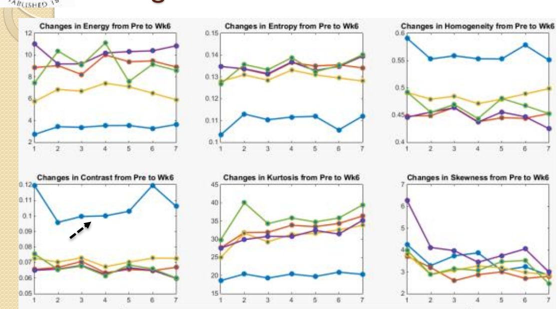


H.Veeraraghavan, N.Tyagi, M. Hunt, N.Riaz, N.Lee, J.O.Deasy. In preparation

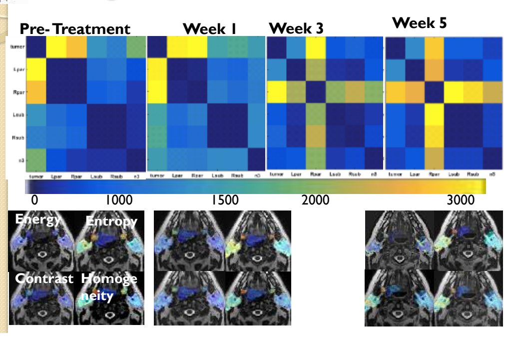
Looking at Longitudinal Trends as Changes in Texture – Patient 1



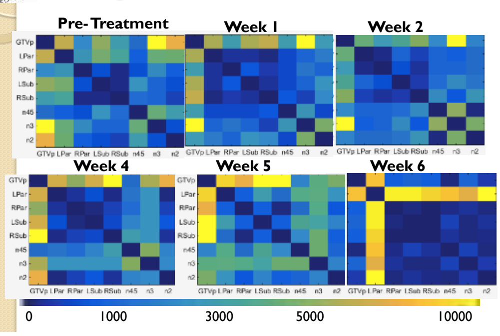
Looking at Longitudinal Trends as Changes in Texture – Patient 3



Longitudinal Trends Inter-Structure Changes over Time – Patient 1



Longitudinal Trends Inter-Structure Changes over Time – Patient 3



Solutions

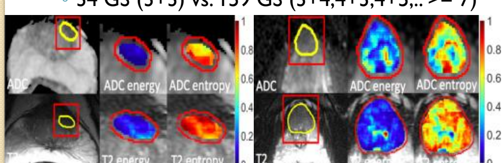
- Highly variable datasets
  - Novel representations of the data
- Highly unbalanced data
  - Sample augmentation-based machine learning
- Segmentation
  - Machine-learning and semi-automatic longitudinal image segmentation methods

Highly Unbalanced Data

- Typical in medical image analysis and radiomics
  - Too many examples from one class (normal pixels) vs. too few (cancer pixels)
  - Too many examples (highly aggressive cancers) vs. too few (benign cancers)

Problems When Classifying with Unbalanced Data

- Classify prostate Gleason scores from MRI
  - 34 GS (3+3) vs. 159 GS (3+4,4+3,4+5,.. >= 7)



GS (3+3 = 6) **GS (4+5 = 9)**

D. Fehr, H.Veeraraghavan, A. Wibmer, T.Gondo, K.Matsumoto, H.A.Vargas, E.Sala, H.Hricak, J.O.Deasy, in review

Classification Results From Different Methods

Method	PZ and TZ Accuracy 34 (3+3=6)vs. 159(>=7)	PZ only Accuracy 23 (3+3=6)vs. 120(>=7)
T-Test SVM	0.83	0.86
RFE-SVM	0.83	0.84
AdaBoost	0.73	0.79
SVM (mADC)	0.82	0.84
SVM (mADC & mT2)	0.82	0.84

Results look surprisingly good regardless of the method used!!

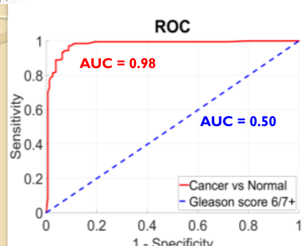
Taking a Closer Look at Results

Youden Index (YI): Specificity + Sensitivity - 1

Method	PZ and TZ:YI 34 (3+3=6)vs. 159(>=7)	PZ only:YI 32 (3+3=6)vs. 120(>=7)
T-Test SVM	0.06	0.24
RFE-SVM	0.03	0.00
AdaBoost	0.11	0.34
SVM (mADC)	0.00	0.00
SVM (mADC & mT2)	0.00	0.00

Results are not looking so good after all!

Taking a Closer Look at Results



Why?
Minority class gets classified as majority class

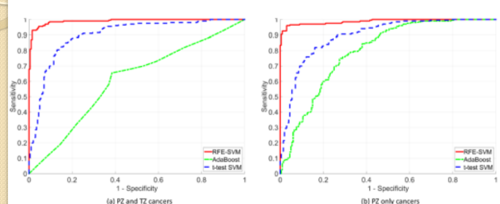
Solution

- Terrible solution:
  - Under sample majority class to the same proportion as the minority class
  - We end up having nothing and over fitting the model

Solution

- Better solution:
  - Oversample the minority class so its similarly represented as the majority class
    - We generate "new" samples in the vicinity of the original samples and thereby help the classifier to model both minority and majority class

Results with Minority Oversampling



Looking at Numbers

Method	PZ and TZ Accuracy (YI) 34 (3+3=6)vs. 159(>=7)	PZ only Accuracy(YI) 23 (3+3=6)vs. 120(>=7)
T-Test SVM	0.84(0.68)	0.74(0.49)
RFE-SVM	0.94(0.92)	0.93(0.86)
AdaBoost	0.64(0.28)	0.72(0.44)
SVM (mADC)	0.61(0.23)	0.65(0.30)
SVM (mADC & mT2)	0.68(0.37)	0.67(0.34)

Results of every classifier improves!!

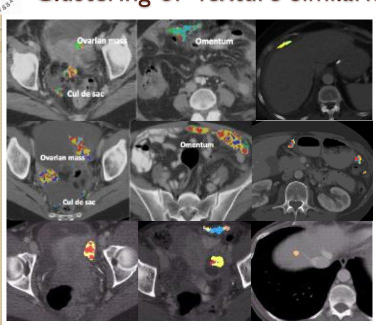
Solutions

- Highly variable datasets
  - Novel representations of the data
- Highly unbalanced data
  - Sample augmentation-based machine learning
- Segmentation
  - Machine-learning and semi-automatic longitudinal image segmentation methods

How about Representing Metastatic Disease Heterogeneity?

- Example: High grade serous ovarian cancers (HGSOC)
  - Patients almost always present with metastatic disease
  - Extent of metastatic spread is highly variable
  - Problem: How do we correlate patients with different extent of disease to outcomes?

Metastatic Site Heterogeneity through Clustering of Texture Similarities



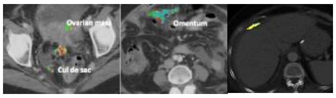
P1- Mesenchymal subtype
Alive: 10.5mo

P2- Differentiated subtype
Alive: 70.4mo

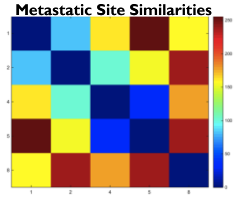
P3- Proliferative subtype
Alive

H.Veeraraghavan, H.A.Vargas, S.Nougaret, J.O.Deasy, H.Hricak, A.S-Charen, E.Sala, in preparation

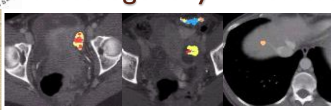
Capturing Metastatic Site Heterogeneity - Mesenchymal



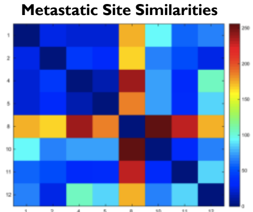
Metastatic Site Similarities



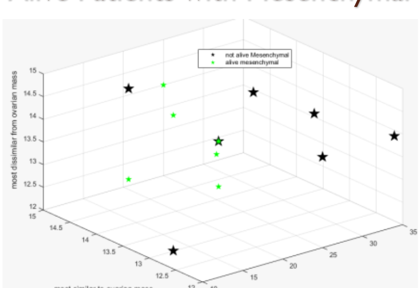
Capturing Metastatic Site Heterogeneity - Proliferative



Metastatic Site Similarities

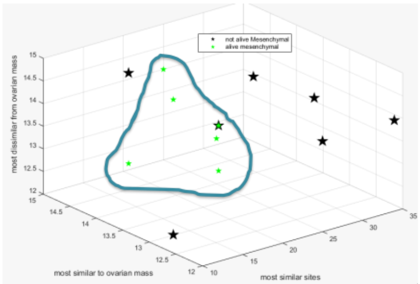


Differences Between Alive vs. Not Alive Patients with Mesenchymal



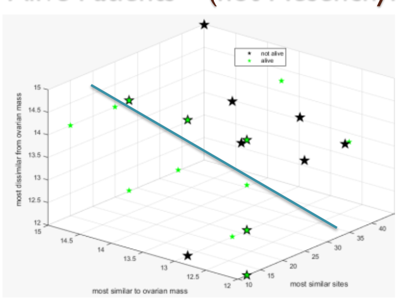
H.Veeraraghavan, H.A.Vargas, S.Nougaret, J.O.Deasy, H.Hricak, A.S-Charen, E.Sala, in preparation

Differences Between Alive vs. Not Alive Patients with Mesenchymal



H.Veeraraghavan, H.A.Vargas, S.Nougaret, J.O.Deasy, H.Hricak, A.S-Charen, E.Sala, in preparation

Differences Between Alive vs. Not Alive Patients – (not Mesenchymal)



H.Veeraraghavan, H.A.Vargas, S.Nougaret, J.O.Deasy, H.Hricak, A.S-Charen, E.Sala, in preparation

Summary of Metastatic Disease Characteristics

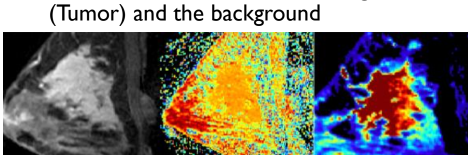
- Patients with good outcomes (survival) tend to have:
  - Most texturally similar sites tend to be like the ovarian mass or cul de sac regardless of disease sub-type
- Patients with poor outcomes (survival) tend to have:
  - Distant metastatic sites tend to be most texturally similar to each other

Conclusions

- Longitudinal radiomics analysis has many challenges
- Some solutions to tackle these challenges are:
  - Extracting appropriate data representation
  - Dealing with unbalanced data
  - Last but not least: Automating volumetric segmentations is important for consistent analysis

How Does it Work?

- Combine multi-parametric MRI (T1pre, T1post1, T1post2, T1post3) and computed image features
- GMM model is a multi-parametric model that extracts a model of the foreground (Tumor) and the background



Source image T2-w post contrast t=1 | Temporal tensor anisotropy image | Temporal tensor trace image