Multi-Atlas Intensity and Label Fusion with Supervised Segmentation Refinement for the Parcellation of Hippocampal Subfields

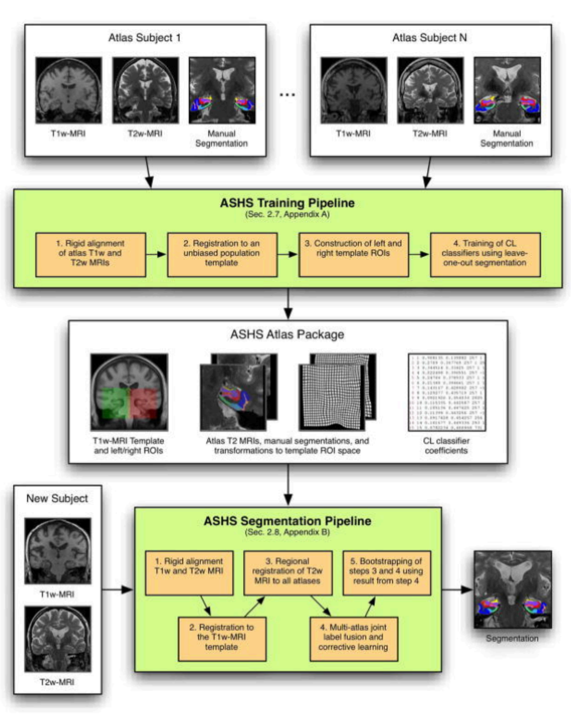
YASSALAB
Translational Neurobiology Laboratory

Nicholas Tustison[1,2], Brian Avants[3], Hongzhi Wang[4] & Michael Yassa[2]

[1]Department of Radiology and Medical Imaging, University of Virginia, [2]Department of Neurobiology and Behavior, University of California, Irvine, [3]Biogen, Cambridge, [4]IBM Research, San Jose

BACKGROUND

- To better characterize progression of neurodegenerative diseases involving the hippocampus (e.g. Alzheimer's disease), significant research focus has been devoted to the accurate delineation of the hippocampus and its subfields.
- Due to the insights gained from "big data" efforts, such as ADNI, automation of subfield hippocampal segmentation techniques is an absolute necessity.
- The well-known Automated Segmentation of Hippocampal Subfields (ASHS)[1] framework has excellent performance and is publicly available.



Training steps
1. atlas registration
2. joint label fusion
3. AdaBoost training

Segmentation steps
1. atlas registration
2. joint label fusion
3. corrective learning
4. repeat steps 1, 2, 3
5. heuristics

- We provide several enhancements/innovations for improved performance in addition to making it publicly available.

METHODS & MATERIALS

Enhancements

Registration
- ANTS → antsRegistration
- B-spline SyN (“-t BSplineSyN[…]”)
- WarpImageMultiTransform → antsApplyTransforms
- Generic label interpolation (“-t GenericLabel[Linear]”)

Joint fusion[2]
- Non-negative least squares
- Multi-threaded
- ITK implementation
- Joint *intensity* fusion*

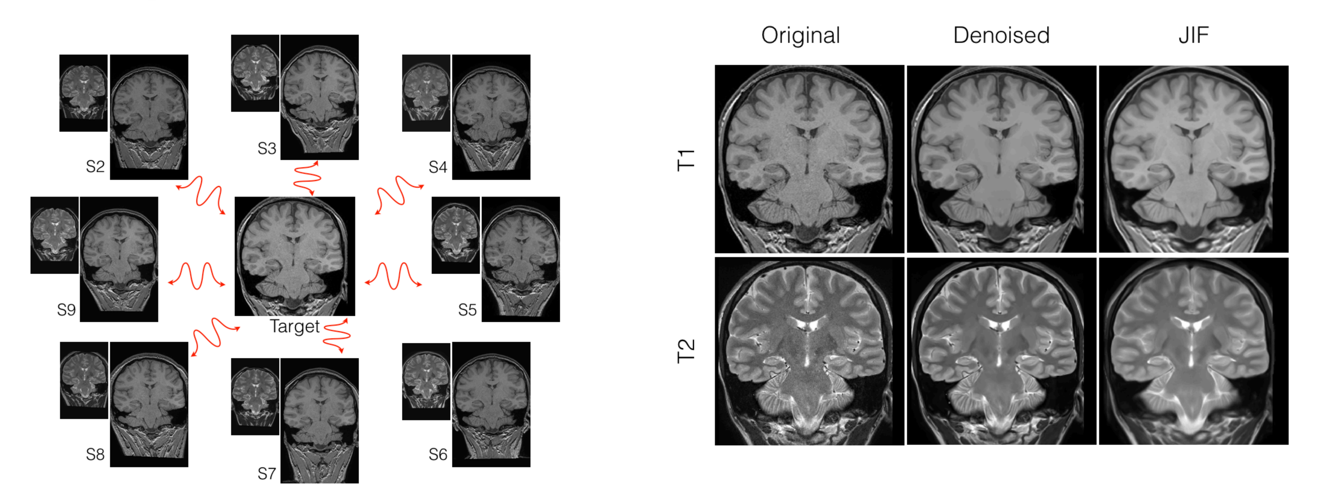
Corrective learning
- Random forests and extreme gradient boosting*
- ANTsR implementation
- Prior knowledge: 2-classes vs. 4-classes*

Evaluation data

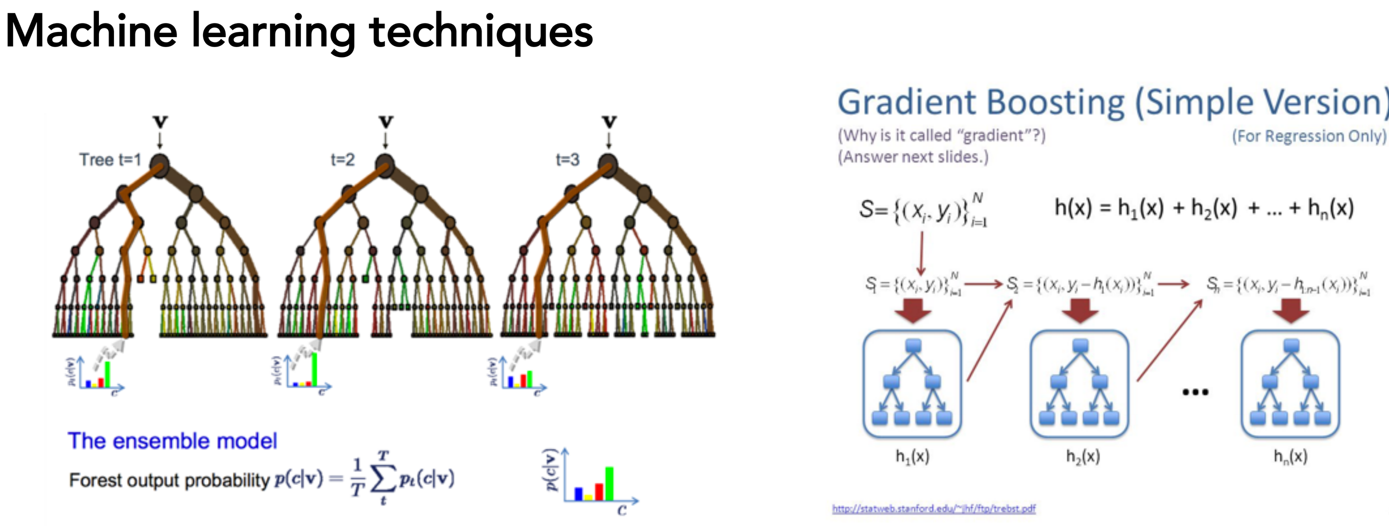
UPenn data[1]
- 29 subjects
- 10 labels per hemisphere
- Hippocampal subfields with extra-hippocampal structures (ERC, PRC, PHC)

UCI data
- 19 subjects
- 3 labels per hemisphere

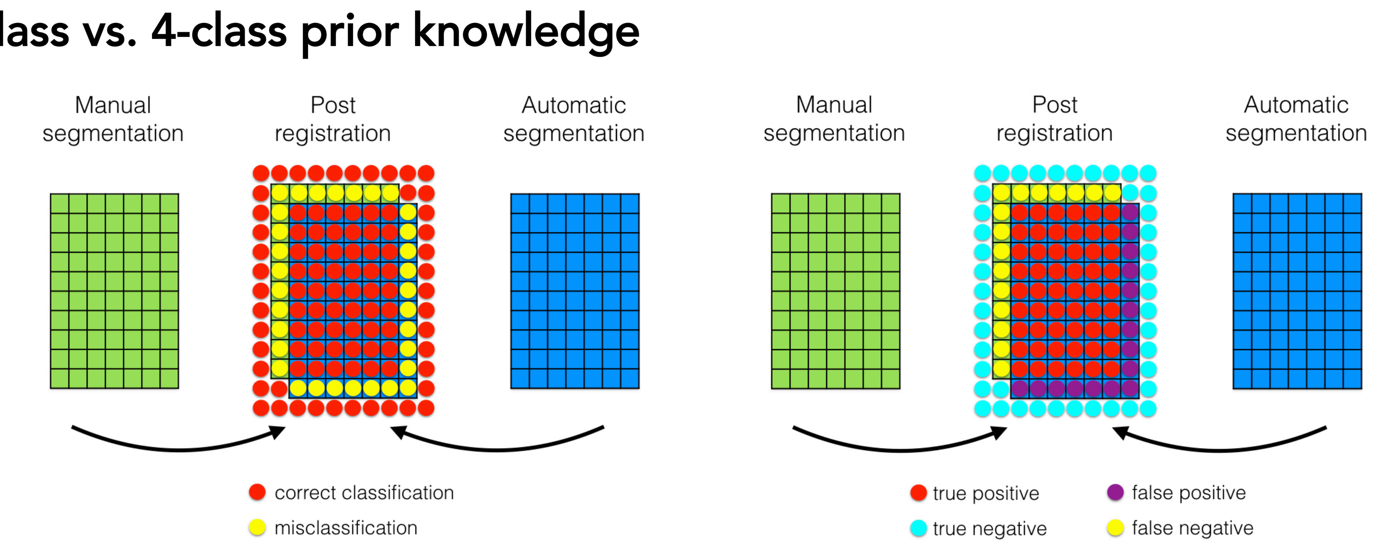
Joint intensity fusion



Machine learning techniques



2-class vs. 4-class prior knowledge



RESULTS

Evaluation: UPenn data

```
> aovLeft <- aov( Dice ~ Method + Label, allDataFrameLeft )
> summary( aovLeft )

              Df Sum Sq Mean Sq F value Pr(>F)
Method         3  0.357  0.1190   20.29 8.87e-13 ***
Label          8 12.809  1.6011  273.04  < 2e-16 ***
Residuals   1032  6.052  0.0059
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> TukeyHSD( aovLeft, "Method" )

  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Dice ~ Method + Label, data = allDataFrameLeft)

$Method
                      diff         lwr        upr     p adj
SVD-NNLS      -0.006890317 -0.02413941 0.01035878 0.7331132
AdaBoost-NNLS  0.030424709  0.01317561 0.04767380 0.0000374
Xgb-NNLS       0.035627356  0.01837826 0.05287645 0.0000008
AdaBoost-SVD   0.037315026  0.02006593 0.05456412 0.0000002
Xgb-SVD        0.042517673  0.02526858 0.05976677 0.0000000
Xgb-AdaBoost   0.005202648 -0.01204645 0.02245174 0.8652726
```

```
> aovRight <- aov( Dice ~ Method + Label, allDataFrameRight )
> summary( aovRight )

             Df Sum Sq Mean Sq F value Pr(>F)
Method        4 0.00458 0.00115   1.772  0.134
Label         3 0.14848 0.04949  76.595 <2e-16 ***
Residuals   372 0.24037 0.00065
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> TukeyHSD( aovRight, "Method" )

  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Dice ~ Method + Label, data = allDataFrameRight)

$Method
                      diff         lwr        upr     p adj
SVD-NNLS      -0.006271697 -0.02404952 0.01150612 0.8007052
AdaBoost-NNLS  0.029442103  0.01166428 0.04721992 0.0001301
Xgb-NNLS       0.034971774  0.01719395 0.05274959 0.0000029
AdaBoost-SVD   0.035713801  0.01793598 0.05349162 0.0000017
Xgb-SVD        0.041243471  0.02346565 0.05902129 0.0000000
Xgb-AdaBoost   0.005529670 -0.01224815 0.02330749 0.8542476
```

Evaluation: UCI data

```
> aovLeft <- aov( Dice ~ Method + Label, allDataFrameLeft )
> summary( aovLeft )

             Df Sum Sq Mean Sq F value Pr(>F)
Method        4 0.0142 0.00354   3.175 0.0139 *
Label         3 0.1113 0.03711  33.265 <2e-16 ***
Residuals   372 0.4150 0.00112
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> TukeyHSD( aovLeft, "Method" )

  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Dice ~ Method + Label, data = allDataFrameLeft)

$Method
                       diff           lwr        upr     p adj
SVD-NNLS       0.0001321447 -1.472049e-02 0.01498478 0.9999999
AdaBoost-NNLS  0.0102401711 -4.612460e-03 0.02509280 0.3245300
Xgb-NNLS       0.0149043158  5.168453e-05 0.02975695 0.0487253
Rf-NNLS        0.0112067368 -3.645894e-03 0.02605937 0.2361193
AdaBoost-SVD   0.0101080263 -4.744605e-03 0.02496066 0.3378764
Xgb-SVD        0.0147721711 -8.046021e-05 0.02962480 0.0520404
Rf-SVD         0.0110745921 -3.778039e-03 0.02592722 0.2472042
Xgb-AdaBoost   0.0046641447 -1.018849e-02 0.01951678 0.9109412
Rf-AdaBoost    0.0009665658 -1.388607e-02 0.01581920 0.9997728
Rf-Xgb        -0.0036975789 -1.855021e-02 0.01115505 0.9601853
```

```
> aovRight <- aov( Dice ~ Method + Label, allDataFrameRight )
> summary( aovRight )

             Df Sum Sq Mean Sq F value Pr(>F)
Method        4 0.00458 0.00115   1.772  0.134
Label         3 0.14848 0.04949  76.595 <2e-16 ***
Residuals   372 0.24037 0.00065
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> TukeyHSD( aovRight, "Method" )

  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Dice ~ Method + Label, data = allDataFrameRight)

$Method
                       diff          lwr         upr     p adj
SVD-NNLS       0.0005491711 -0.010754534 0.011852877 0.9999289
AdaBoost-NNLS  0.0054772368 -0.005826469 0.016780942 0.6737533
Xgb-NNLS       0.0091173289 -0.002186377 0.020421034 0.1778540
Rf-NNLS        0.0060720658 -0.005231640 0.017375771 0.5810319
AdaBoost-SVD   0.0049280658 -0.006375640 0.016231771 0.7542446
Xgb-SVD        0.0085681579 -0.002735548 0.019871863 0.2318887
Rf-SVD         0.0055228947 -0.005780811 0.016826600 0.6667792
Xgb-AdaBoost   0.0036400921 -0.007663613 0.014943798 0.9031737
Rf-AdaBoost    0.0005948289 -0.010708877 0.011898534 0.9999023
Rf-Xgb        -0.0030452632 -0.014348969 0.008258442 0.9473355e
```

DISCUSSION

Conclusions

- We provide an open-source pipeline for segmentation of hippocampal subfields using consensus labeling and refinement, which outperforms existing software.
- Results were applied to T1/T2 data but the framework is sufficiently general to accommodate other imaging protocols (e.g. T1-only).
- Advanced machine learning techniques (random forests and extreme gradient boosting) were explored providing unique performance characteristics.

Future directions

- Develop and implement hippocampal-specific feature images using established manual protocols.
- Integrate into Advanced Normalization Tools longitudinal cortical thickness pipeline[3] for EC-specific cortical thickness estimation.

REFERENCES

1. Yushkevich et al. (2015). Automated Volumetry and Regional Thickness Analysis of Hippocampal Subfields and Medial Temporal Cortical Structures in Mild Cognitive Impairment. *Human Brain Mapping.*
2. Tustison et al. (2017). A patch-based framework for new ITK functionality: Joint fusion, denoising, and non-local super-resolution. *Insight Journal.*
3. Tustison et al. (2017). The ANTs longitudinal cortical thickness pipeline. ADPD 2017.

UCI MIND Institute for Memory Impairments and Neurological Disorders

UCI CNLM Ayala School of Biological Sciences | Center for the Neurobiology of Learning & Memory

Funding support provided by NIMH R01 MH102392 and NIA R21 AG049220, P50 AG16573.

Contact: ntustiso@uci.edu
www.yassalab.org