Paper #90

Segmentation of Brain Tissues from Infant MRI Records Using Machine Learning Techniques

Béla Surányi Levente Kovács László Szilágyi

Sapientia University of Transylvania, Romania

Óbuda University, Budapest, Hungary

Motivation

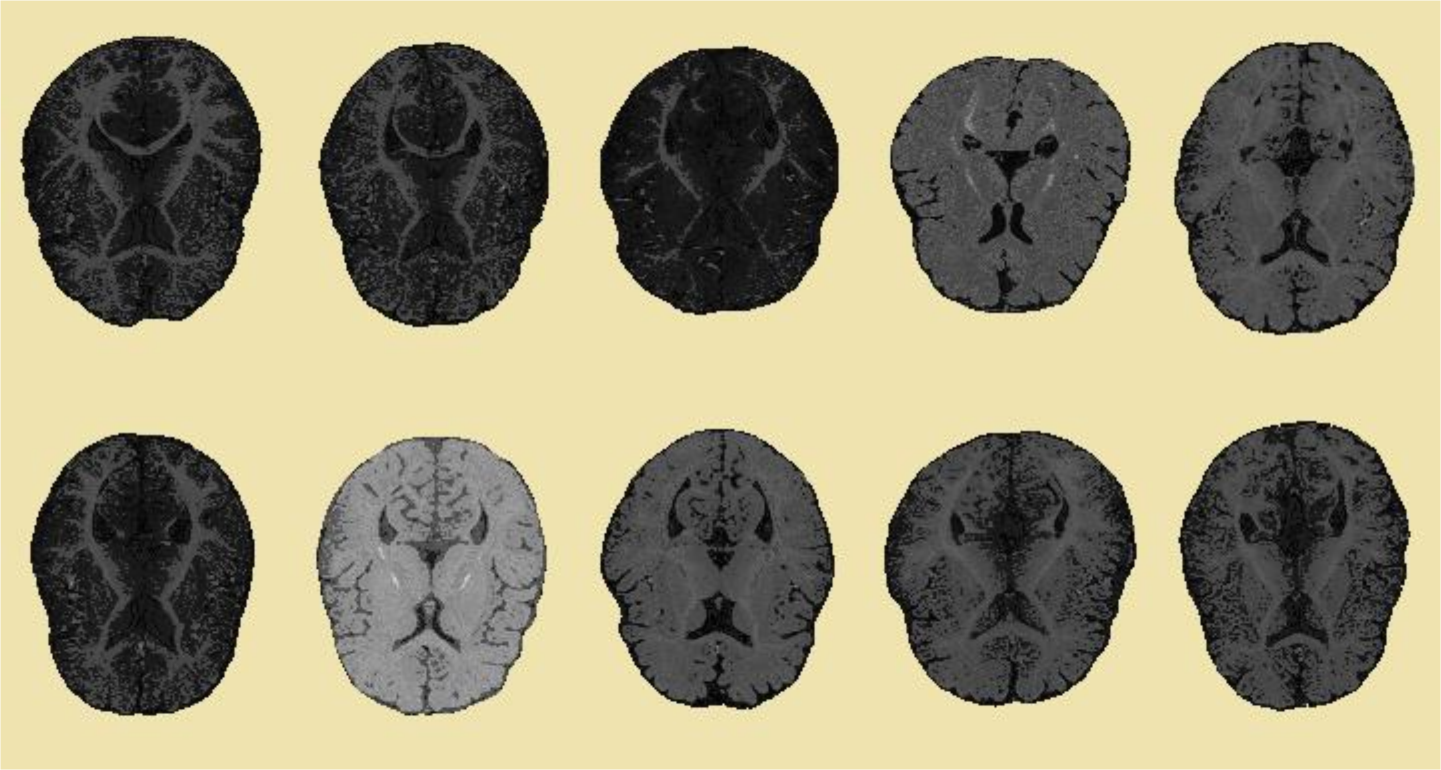
- Large-scale research to study the correlation between
  - The development of the infant brain (first 1-2 years)
  - And the occurrence of various diseases in later years
- Why is it useful?
  - To predict the diseases that are likely to affect babies in the future
- Main goal
  - Accurately separate the three main tissue types (WM, GM, CSF) in the infant brain based on T1 and T2 weighted volumetric MRI
- Main difficulty
  - WM and GM have almost the same appearance at the age of 6 month

Input Data

- iSeg Challenges, years 2017 and 2019
- iSeg-2017 train dataset
  - 10 volumes
- Multispectral (T1, T2), T2 volume registered to T1
- Slices of 144 x 192 pixels, 100-110 slices per volume
- Each pixel represents 1mm^3 of tissues
- Skull removed
- Ground truth (GT):
  - white matter (WM), grey matter (GM), cerebro-spinal fluid (CSF), external pixel
- 700k to 900k brain pixels per volume

Difficulties

- Major
  - Histograms need normalization
  - Intensity inhomogeneity
- Minor
  - Missing data



Procedure

- Preprocessing
  - Histogram normalization, using the classical method of Nyúl et al. (2000)
  - Feature generation
    - 2 observed features (T1, T2), 16 morphological features, 3 relative coordinates
- Classification
  - Six different machine learning methods involved
    - k nearest neighbors, AdaBoost, Random forest, multi-layer perceptron, decision tree, logistic regression
  - No post processing, we wanted to compare the direct outcome of classification
- Statistical evaluation
  - Accuracy indicators for separate tissue types, and global accuracy

Feature generation

- Feature vector contains 21 features
- 2 observed features (T1, T2)
- Averaged values of T1 and T2
  - 3x3, 5x5, 7x7, 9x9, 11x11 neigborhoods (only brain pixels)
- Average minimum and maximum of T1 and T2
  - Spatial 3x3x3 neighborhood (only brain pixels)
- Relative coordinates of the pixel: x, y, z
- All feature values between 1...255, as 0 is reserved for external pixels

Machine learning methods

- In all cases, 9 volumes are train data and 1 is test data
- All volumes take turns to serve as test data
- Algorithms: OpenCV implementation (ver. 3.x)
- kNN: 10k train pixels from each volume, k=13
- Decision tree: 500k train pixels/volume, maximum depth = 18
- Random forest: 100k train pixels/volume, max. 45 trees of maximum depth = 26
- Multi-layer perceptron: two hidden layers of sizes 32 and 16, 10k train pixels/volume
- AdaBoost: 30k train pixels/volume
- Logistic regression: 5k train pixels/volume
- All methods configured to perfom as fine as possible in 5 minutes (training and testing), except LogReg which performed much faster (<15 sec)

Measuring accuracy

- Statistical accuracy indicators based on the confusion matrix
- $C = (c_{ij})$ with $i, j \in \{CSF, GM, WM\} = \Lambda$
- c_{ij} = number of pixels with GT= i and label = j

- Sum of row i: $\rho_i = \sum_{j \in \Lambda} c_{ij}$
- Sum of column j: $\kappa_j = \sum_{i \in \Lambda} c_{ij}$
- Sum of diagonal elements: $\delta = \sum_{i \in \Lambda} c_{ii}$
- Sum of all elements: $\sigma = \sum_{i \in \Lambda} \rho_i = \sum_{j \in \Lambda} \kappa_j$

- Sensitivity (or true positive rate or recall) with respect to class $i \in \Lambda$, defined as

$$\mathrm{TPR}_i = c_{ii}/\rho_i \quad ,$$

- Positive predictive value (or precision) with respect to class $j \in \Lambda$, defined as

$$\mathrm{PPV}_j = c_{jj}/\kappa_j \quad ,$$

- Specificity (or true negative rate) with respect to class $i \in \Lambda$, defined as

$$\mathrm{TNR}_i = \frac{\sigma - \rho_i - \kappa_i + c_{ii}}{\sigma - \rho_i} \quad ,$$

- Dice similarity coefficient (or Dice score or F1-score) of class $i \in \Lambda$, defined as

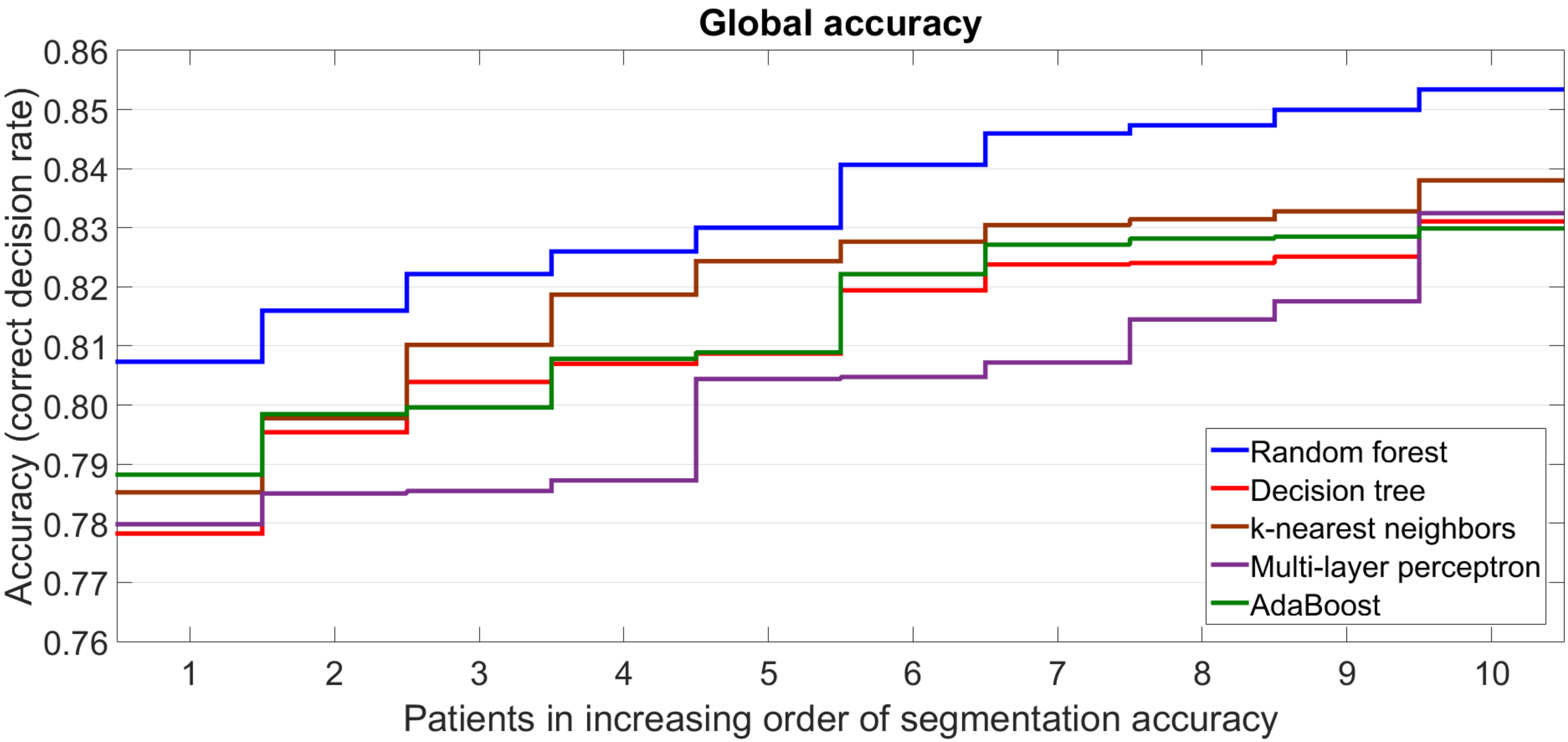
$$\mathrm{DSC}_i = \frac{2 \times c_{ii}}{\rho_i + \kappa_i} \quad ,$$

- and overall accuracy defined as $\mathrm{ACC} = \delta/\sigma$.

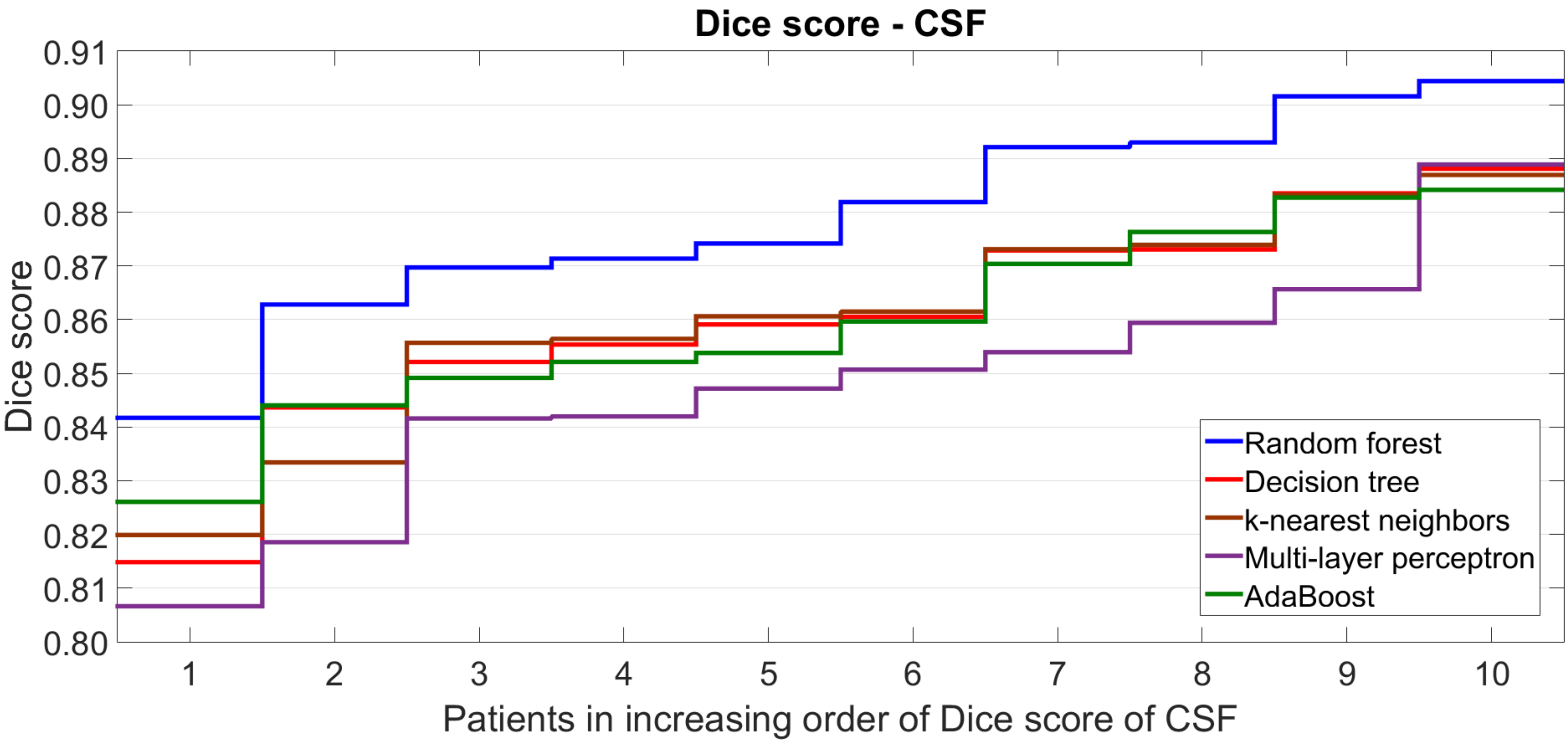
Global accuracy indicators for various tissue types

Classifier algorithm	Dice score			Sensitivity (TPR)			Precision (PPV)			Overall accuracy
	CSF	GM	WM	CSF	GM	WM	CSF	GM	WM	
Random forest	0.8793	0.8357	0.7963	0.8629	0.8779	0.7530	0.9007	0.7997	0.8491	0.8339
k-nearest neighbors	0.8605	0.8208	0.7881	0.8267	0.8560	0.7660	0.9014	0.7909	0.8156	0.8197
AdaBoost	0.8598	0.8132	0.7809	0.8435	0.8401	0.7604	0.8818	0.7904	0.8069	0.8139
Decision tree	0.8603	0.8139	0.7718	0.8350	0.8524	0.7400	0.8922	0.7809	0.8104	0.8117
Multi-layer perceptron	0.8474	0.8043	0.7619	0.8282	0.8352	0.7349	0.8749	0.7773	0.7952	0.8019
Logistic regression	0.7999	0.7299	0.6809	0.7624	0.7584	0.6738	0.8515	0.7051	0.6996	0.7310

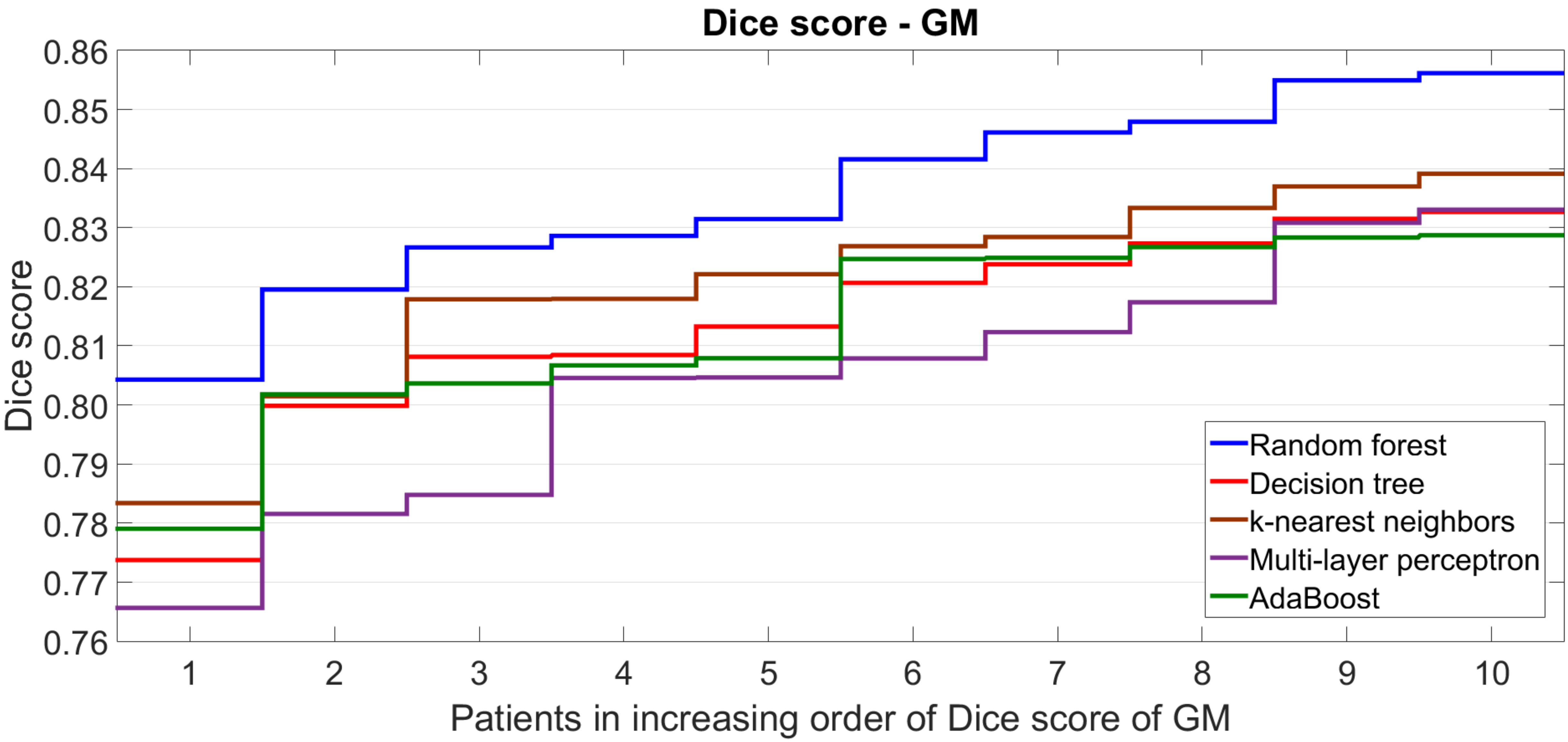
Rate of correct decision – all tissues



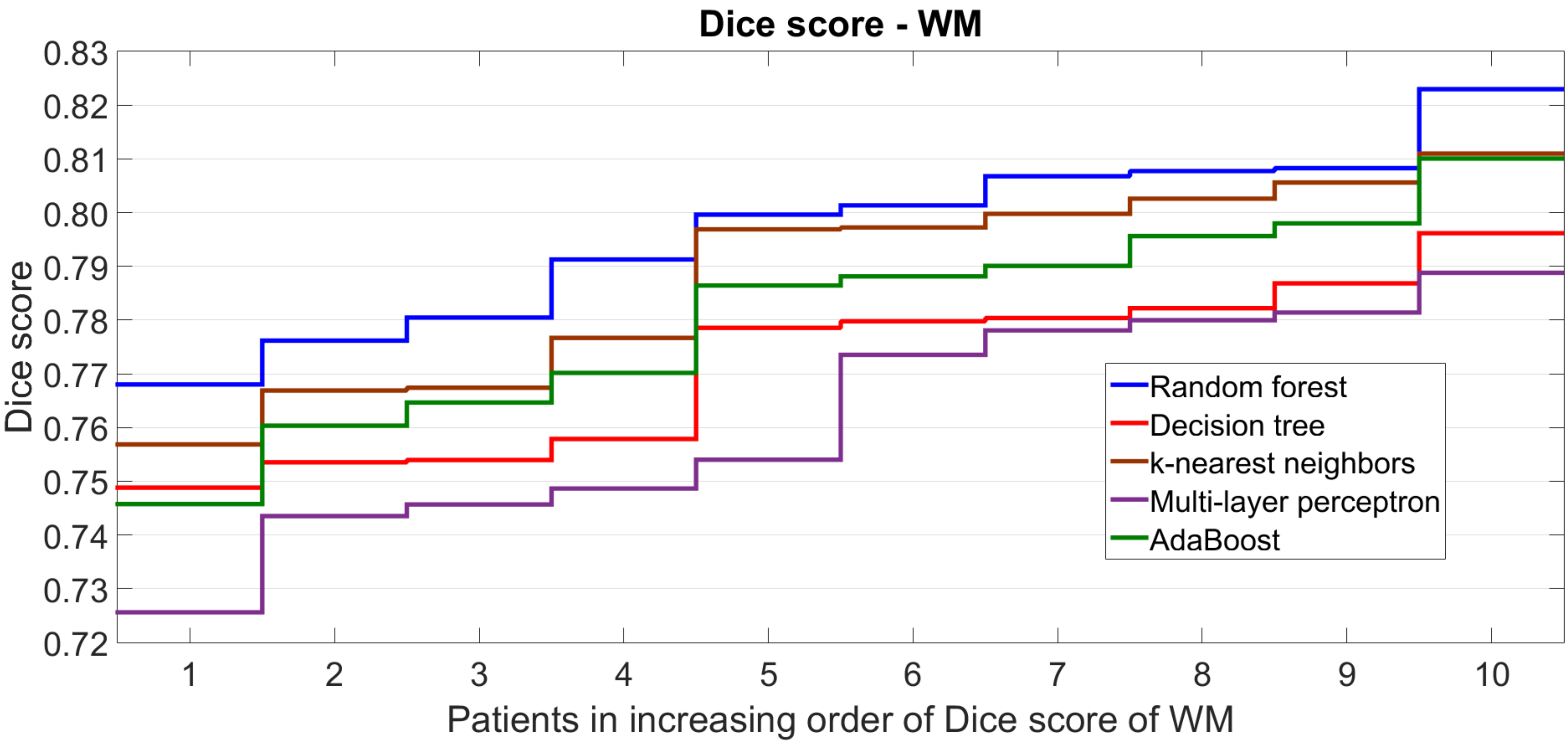
Rate of correct decision – CSF



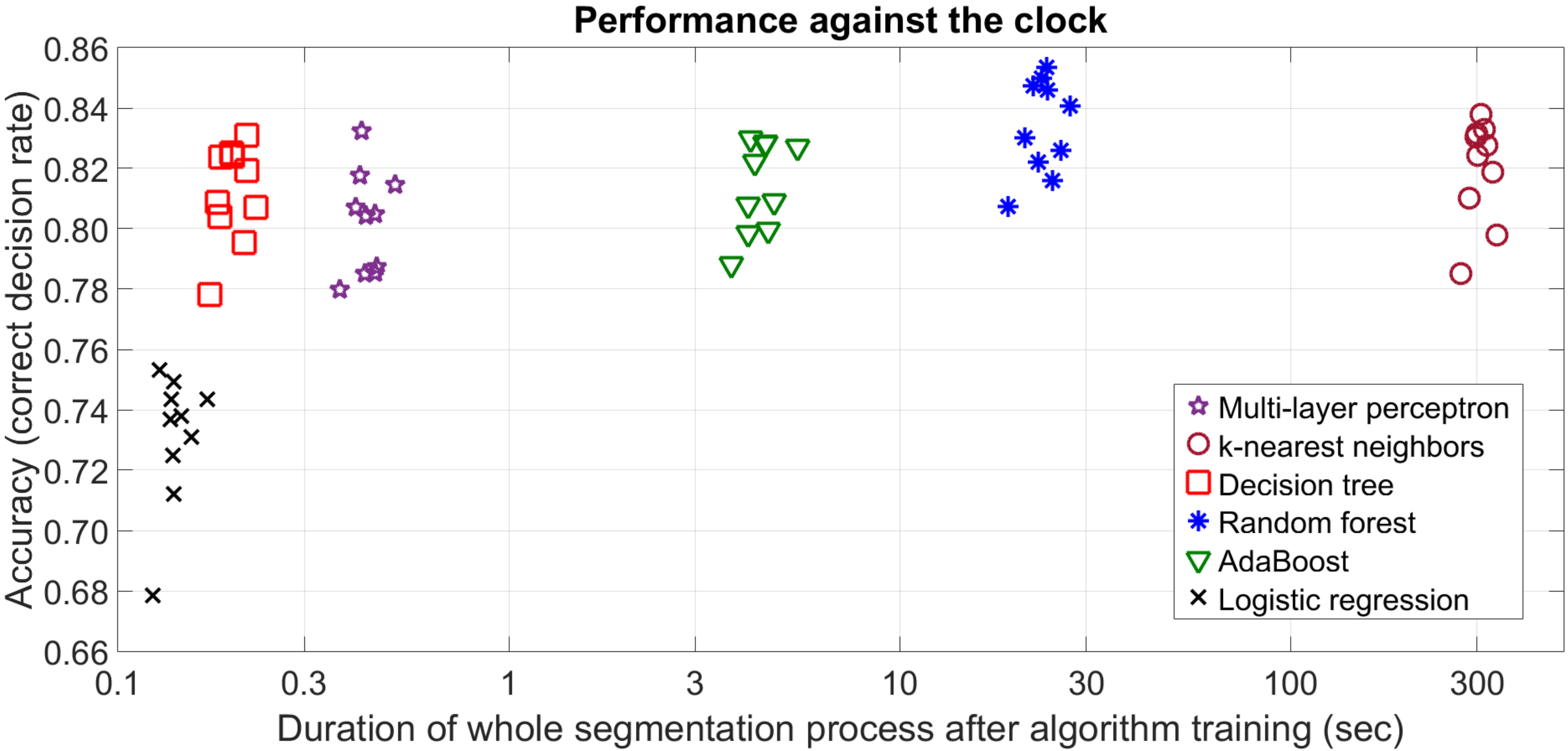
Rate of correct decision – grey matter



Rate of correct decision – white matter



Efficiency

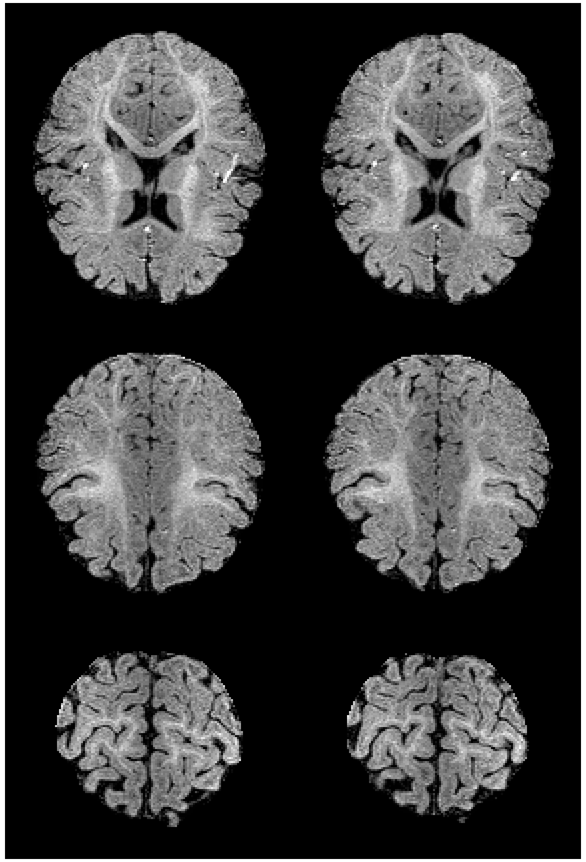

Performance against the clock
Accuracy (correct decision rate)
0.86
0.84
0.82
0.80
0.78
0.76
0.74
0.72
0.70
0.68
0.66
0.1
0.3
1
3
10
30
100
300
Duration of whole segmentation process after algorithm training (sec)
Multi-layer perceptron
k-nearest neighbors
Decision tree
Random forest
AdaBoost
Logistic regression

Accuracy Ranking

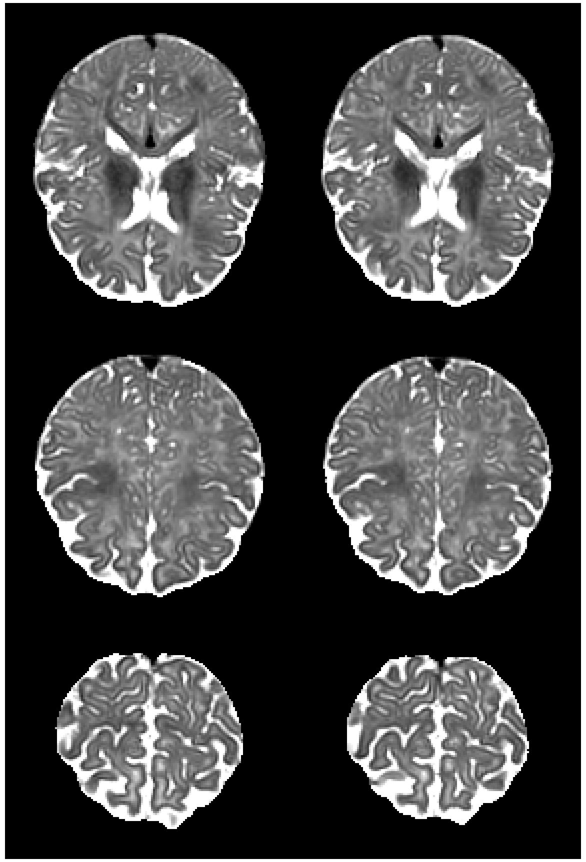
- "Olympic Games" of infant brain tissue segmentation
  - 1st, 2nd, 3rd places
    - for each MRI record
    - for each indicator and each tissue class

Classifier algorithm	ACC			Dice score			Recall (TPR)			Precision (PPV)			Total		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
Random forest	10	0	0	29	1	0	14	9	3	25	3	2	78	13	5
k-nearest neighbors	0	8	2	1	21	4	7	6	6	2	17	4	10	52	16
Multi-layer perceptron	0	0	1	0	2	1	8	4	2	3	6	2	11	12	6
AdaBoost	0	2	4	0	5	14	1	8	10	0	2	11	1	17	39
Decision tree	0	0	3	0	1	11	0	3	9	0	2	11	0	6	34
Logistic regression	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

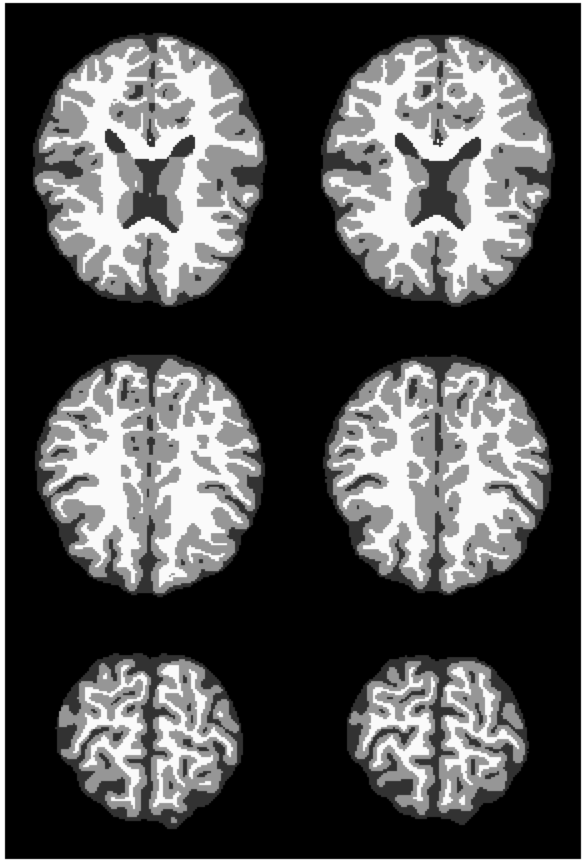
Segmentation result



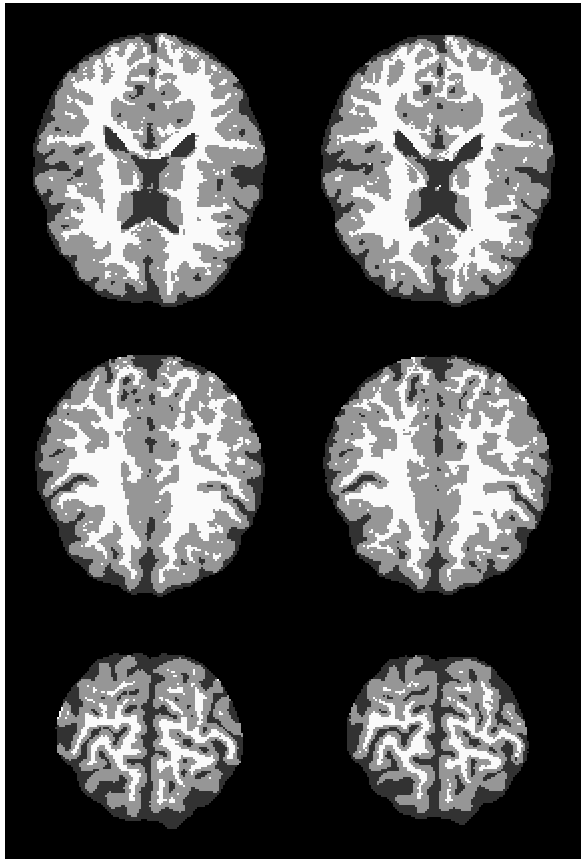
T1



T2



Ground truth



Result

Conclusions

- As preliminary result, the achieved accuracy is promising
- Fine tuning and post processing will improve accuracy

- Future:
  - CNN + deep learning methods

Acknowledgement

- Sapientia Institute for Research Programs
- European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program (grant agreement No 679681)
- Hungarian Academy of Sciences through the János Bolyai Fellowship program
- ÚNKP 20-5 New National Excellence Program of the Ministry of Human Capacities of Hungary