# Title of the Paper

Michael Mueller, Jan Riedo, Michael Rebsamen

***Abstract*—Bla**

***Index Terms*—MRI, Segmentation, Machine Learning, DF, kNN, SVM**

## I. Introduction

Segmentation of brain tissues from magnetic resonance images (MRI) has many clinical applications. Clinicians gain useful information from a separation of tissue into its three main anatomical types: white matter, grey matter, and ventricles. However, manual segmentation of MRI is a labour-intensive task requiring expert skills. Fully automatic approaches for brain tissue segmentation are therefore a topic of active research. A good algorithm classifies the tissue types with high accuracy across a variety of images from different patients. Such a classification is a typical task for machine learning. These algorithms tend to perform well given enough training data during the learning phase. The availability of ground-truth data in sufficient quantity and quality for supervised learning is a particular challenge when working with medical images due to privacy concerns and the costs for manual segmentation. Optimization of the learning phase with a limited number of training data is therefore required.

FIXME: kNN is a popular classification method for MR data and has successfully been applied in MR brain segmentation [1]–[3]

## II. Methods

### A. Dataset

All experiments were conducted on a subset of 100 unrelated subjects from a dataset provided by the *Human Connectome Project* [4]. From each individual, a total of eight 3-tesla head MRI are available: T1 and T2-weighted image volumes not skull-stripped (but defaced for anonymization) and skull-stripped with a bias field correction, and both modalities once in native T1 space and once in MNI-atlas space [5].

Ground-truth labels are automatically generated using *FreeSurf*, assigning each voxel either to background, white matter, grey matter, or ventricles. The dataset was split in a training set with 70 images and a test set with 30 images.

### B. Pipeline

TODO: Describe whole pipeline (registration, preprocessing, feature extraction, ML classification, postprocessing, evaluation

Biomedical Engineering, University of Bern
Authors e-Mail: michael.mueller@students.unibe.ch, jan.riedo@students.unibe.ch, michael.rebsamen@students.unibe.ch

### C. Training

TODO: Describe training of machine learning algorithms
TODO: SVM gridsearch for hyperparameter tuning?

### D. Performance Evaluation

TODO: Describe metric (dice score)

### E. Infrastructure

TODO: Describe UBELIX, libraries

## III. Results

TODO: JR: subplot with 3 result images (good dice bad result, good dice good result, bad ground truth
TODO: JR: DF hyperparameter optimization, 3DPlot The decision forest algorithm was enhanced with normalized features, a higher number of ventricle voxels in the training set and the optimization of the hyperparameters (see Fig. III). With this settings, the max dice coefficient was lifted from 0.703 to 0.754. This result was achieved with 80 trees and 3000 max nodes.

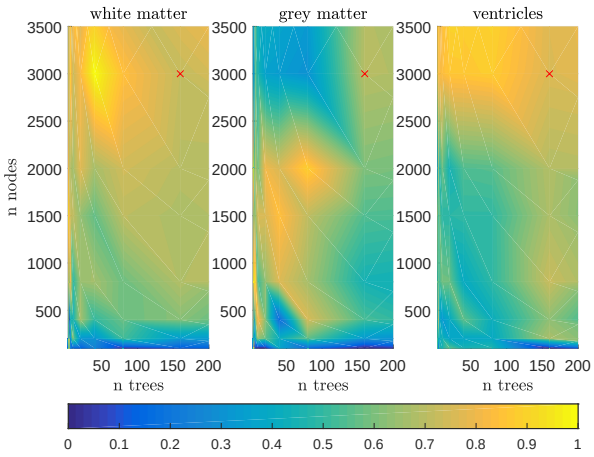


Fig. 1. Plot of grid search for white matter, grey matter and ventricles based on the hyperparameters number of trees and maximum nodes. Color does not represent dice, the data is stretched individually for all three plots.

TODO: JR: kNN optimization

Statistical distribution of the dice coefficients can be seen in Fig. III. DF and SVM achieve a similar mean dice score but SVM has a lower variance for the ventricles.

## IV. Discussion

TODO: challenge with quality of ground truth
TODO: feature importance

TABLE I
PERFORMANCE COMPARISON OF ML ALGORITHMS

| Features | Size Dataset | DF | GMM | kNN | SGD | SVM |
|---|---|---|---|---|---|---|
| All (f1-f7) | 3 | - | - | 0.70/0.57/0.48 | - | - |
| | 12 | 0.84/0.80/0.52 | 0.00/0.78/0.00 | 0.75/0.66/0.67 | - | 0.83/0.81/0.58 |
| | 70 | 0.85/0.80/0.60 | - | 0.79/0.76/0.73 | 0.82/0.80/0.33 | 0.84/0.82/0.61 |
| Coordinates only (f1-f3) | 3 | - | - | 0.70/0.55/0.41 | - | - |
| | 12 | - | - | 0.74/0.63/0.56 | - | - |
| | 70 | - | - | 0.77/0.71/0.62 | - | - |
| All non-coordinates (f4-f7) | 3 | - | - | 0.85/0.80/0.45 | - | - |
| | 12 | - | - | 0.85/0.81/0.45 | - | - |
| | 70 | - | - | 0.85/0.81/0.54 | - | - |

Overview of achieved accuracy for the different algorithms. Mean dice scores for white matter/grey matter/ventricles. f1-f3: Coordinate features, f4: T1 intensity, f5: T1 gradient, f6: T2 intensity, f7: T2 gradient.

TABLE II
RUNTIME

| Features | Size Dataset | DF | GMM | kNN | SGD | SVM |
|---|---|---|---|---|---|---|
| All (f1-f7) | 3 | - | - | 14.4/6564.6 | - | - |
| | 12 | - | - | 39.6/7548.7 | - | - |
| | 70 | - | - | 223.1/8724.6 | - | - |
| Coordinates only (f1-f3) | 3 | - | - | 10.4/4391.5 | - | - |
| | 12 | - | - | 34.7/5449.3 | - | - |
| | 70 | - | - | 196.4/6112.8 | - | - |
| All non-coordinates (f4-f7) | 3 | - | - | 10.1/10084.7 | - | - |
| | 12 | - | - | 34.6/18768.6 | - | - |
| | 70 | - | - | 194.2/16555.7 | - | - |

FIXME: Overview of the computation time in seconds for all algorithms (training time/testing time). Computation time includes pre- and post-processing.

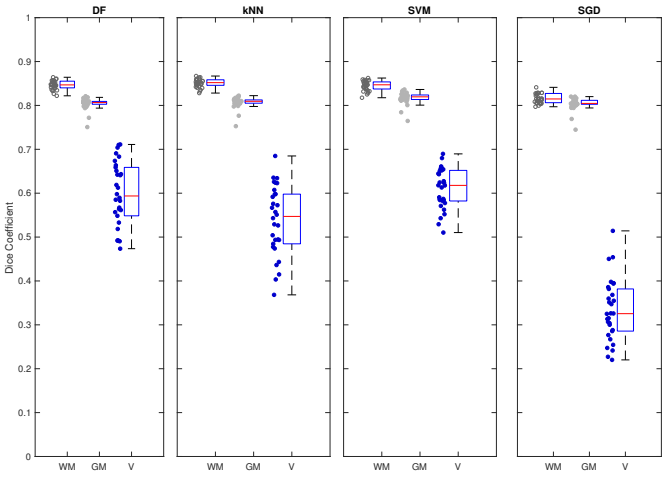


Fig. 2. Distribution of dice coefficients with optimal hyper-parameters for each algorithm on the full training set of 70 images.

## V. CONCLUSION

## ACKNOWLEDGMENT

Calculations were performed on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern.

## REFERENCES

[1] P. Anbeek, K. L. Vincken, M. J. van Osch, R. H. Bisschops, and J. van der Grond, "Probabilistic segmentation of white matter lesions in MR imaging," *NeuroImage*, vol. 21, no. 3, pp. 1037–1044, mar 2004.

[2] C. A. Cocosco, A. P. Zijdenbos, and A. C. Evans, "A fully automatic and robust brain MRI tissue classification method," *Medical Image Analysis*, vol. 7, no. 4, pp. 513–527, dec 2003.

[3] S. Warfield, M. Kaus, F. A. Jolesz, and R. Kikinis, "Adaptive, template moderated, spatially varying statistical classification," *Medical Image Analysis*, vol. 4, no. 1, pp. 43–55, mar 2000.

[4] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium *et al.*, "The wu-minn human connectome project: an overview," *Neuroimage*, vol. 80, pp. 62–79, 2013.

[5] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike *et al.*, "A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (icbm)," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 356, no. 1412, pp. 1293–1322, 2001.