Decision Forests for Tissue-specific Segmentation of High-grade Gliomas in Multi-channel MR

D. Zikic[1], B. Glocker[1], E. Konukoglu[1], A. Criminisi[1], C. Demiralp[2], J. Shotton[1], O.M. Thomas[3,4], T. Das[3], R. Jena[3], S.J. Price[3,5]

[1]Microsoft Research Cambridge
[2]Brown University, Providence, RI, USA
[3]Cambridge University Hospitals, Cambridge, UK
[4]Department of Radiology and [5]Department of Clinical Neurosciences, Cambridge University

Microsoft Research

Goal: Automatic segmentation of tumorous tissue types

Active Cells (**AC**), Necrotic Core (**NC**), and Edema (**E**)

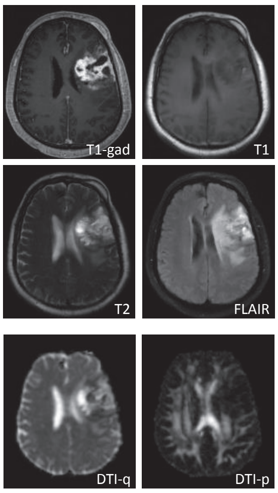
Challenge:

High variability in structure, location, shape and appearance

Motivation:

- **Clinical practice: Time savings**
  Initialize interactive segmentation for treatment
- **Research: Volumetric measurements for individual tissues**
  Quantification of progress/treatment in follow-up and research studies

Our Approach: Context-sensitive Classification Forest



Input data I_{MR} (MRI + DTI)

1. Initial Class Probability Estimates

- GMM-based model of *local* multi-channel intensity

Model-B | Model-AC | Model-NC | Model-E

- Advantage: Speed / Drawback: Quality (noisy results)

Initial probabilities used as additional input channels

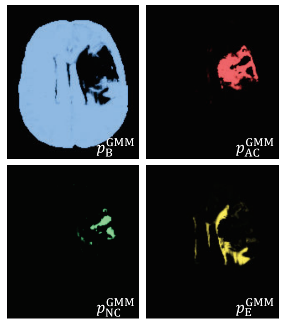
$I = (I_{\mathrm{MR}}, p_c^{\mathrm{GMM}})$

2. Context-sensitive Classification Forest

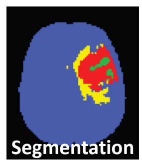
- Context-sensitivity through:
  - Initial class probabilities as additional input channels
  - Non-local, parameterized features (→high-dimensional feature space)

→ Learned application/data-specific regularization effect

- General properties of Classification Forests:
  - Inherent multi-label classification
  - Efficiency for high-dimensional feature spaces
  - Good generalization ability



Initial tissue probabilities p_c^{GMM}



Initial Probabilities: GMM-based posteriors

- Train likelihood model of local, multi-channel MR intensity for each class c (GMM-based) $p(I_{\mathrm{MR}}|c)$
- Probability of class c given intensity at point x (posterior probability)

$$p_c^{\mathrm{GMM}}(x) = p^{\mathrm{GMM}}(c|I_{\mathrm{MR}}(x)) = \frac{p(I_{\mathrm{MR}}(x)|c)\, p_c}{\sum_{c_j} p(I_{\mathrm{MR}}(x)|c_j)\, p_{c_j}}$$

with empirical class probability p_c

Forest Feature Types: non-local, parametric, intensity-based

- **Type 1: Intensity Difference**

$$f^1_{j_1,j_2,v}(x,I) = I_{j_1}(x) - I_{j_2}(x+v)$$

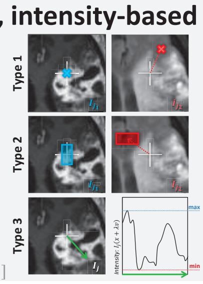
- **Type 2: Intensity Difference of Region Means**

$$f^2_{j_1,j_2,s_1,s_2 v}(x,I) = \mu(N^{s_1}_{j_1}(x)) - \mu(N^{s_2}_{j_2}(x+v))$$

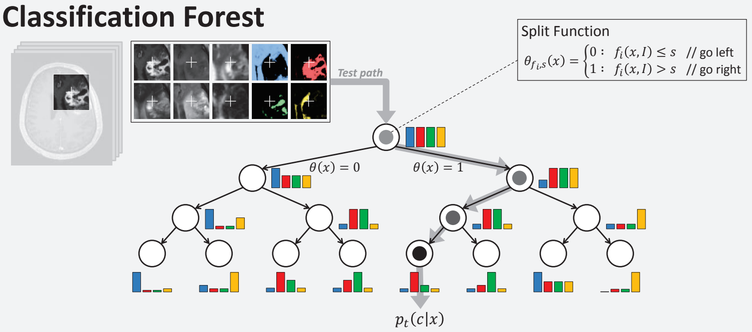
- **Type 3: Intensity Range along Ray**

$$f^3_{j,v}(x,I) = \max_{\lambda} I_j(x+\lambda v) - \min_{\lambda} I_j(x+\lambda v)$$

$v = (v_x, v_y, v_z)$: offset vector; N^s: cuboid with side lengths $s = (s_x, s_y, s_z)$; $\lambda \in [0,1]$



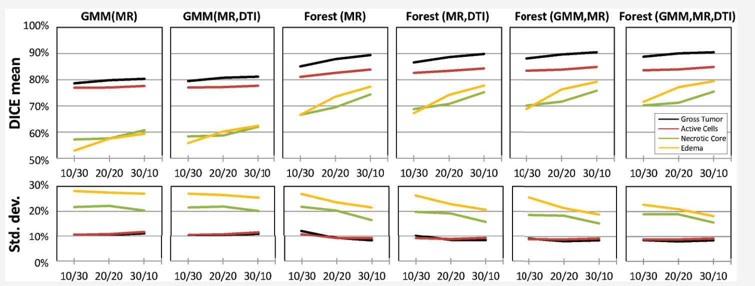
Classification Forest



- **Context-sensitive Features** *(data representation)*
  - Feature vector $f(x,I)_{\mathrm{p}}$ represents a data point at x (spatial point in scan) by non-local, parameterized, and intensity-based features
  - Feature space: $\mathcal{F} = \{\ f^{\mathrm{t}}(x,I)_{\mathrm{p_t}}\ :\ \mathrm{t} \in \mathrm{Type}\,,\ \ \mathrm{p} \in \mathrm{P} \subset \mathbb{Z}^{|\mathrm{p_t}|}\ \}$
    → very high dimension of $\mathcal{F}$: number of unique parameter settings $\mathrm{p_t}$ for all feature types t
  - *N. B.: Automatically learned feature parameters, instead of manual feature design*
- **Training** *(learning from training data with manual label annotations)*
  - Determine split functions $\theta_{f_i,s}$ at nodes (→ tree structure, statistics at leaves): Estimate splitting dimension f_i (from *randomly chosen feature subspace* $\mathcal{F}'$) and split threshold s for f_i, so as to optimize Information Gain ('clustering of classes')
- **Testing** *(determining the label for unseen data point)*
  - For each tree t, apply split tests to reach a leaf, and use the resulting $p_t(c|x)$ in overall prediction

$$p(c|x) = \frac{1}{n}\sum_{t=1}^{n} p_t(c|x)$$

Evaluation



Dataset of High-grade Glioma Patients

- 40 multi-channel MR scans (pre-treatment): T1-gad, T1, T2, FLAIR, DTI-p, DTI-q
- Pre-processing: skull-stripping, affine intra-patient registration, resampling to 2mm, intensity mean alignment per channel (global multiplicative factor)
- Manual segmentations of **AC**, **NC**, **E** in 3D; gross tumor: **GT** = **AC** ∪ **NC**

Experiment Setup

- Leave-N-out with N=10,20,30 → Training/Testing ratios of 10/30, 20/20, 30/10
- Repeated 10x with random test set draws → 600 test segmentations per approach
- Evaluated approaches (each with and without DTI channels):
  - Baseline: Initial probabilities (GMM-based, local intensity only)
  - Forest without initial probabilities
  - Proposed approach: Forest with initial probabilities
- Settings: number of trees n=40, depth d=20, per-node-feature-subspace size $|\mathcal{F}'|$=200
- Timings: training per tree: 10-25 min, testing per image 2-3 min