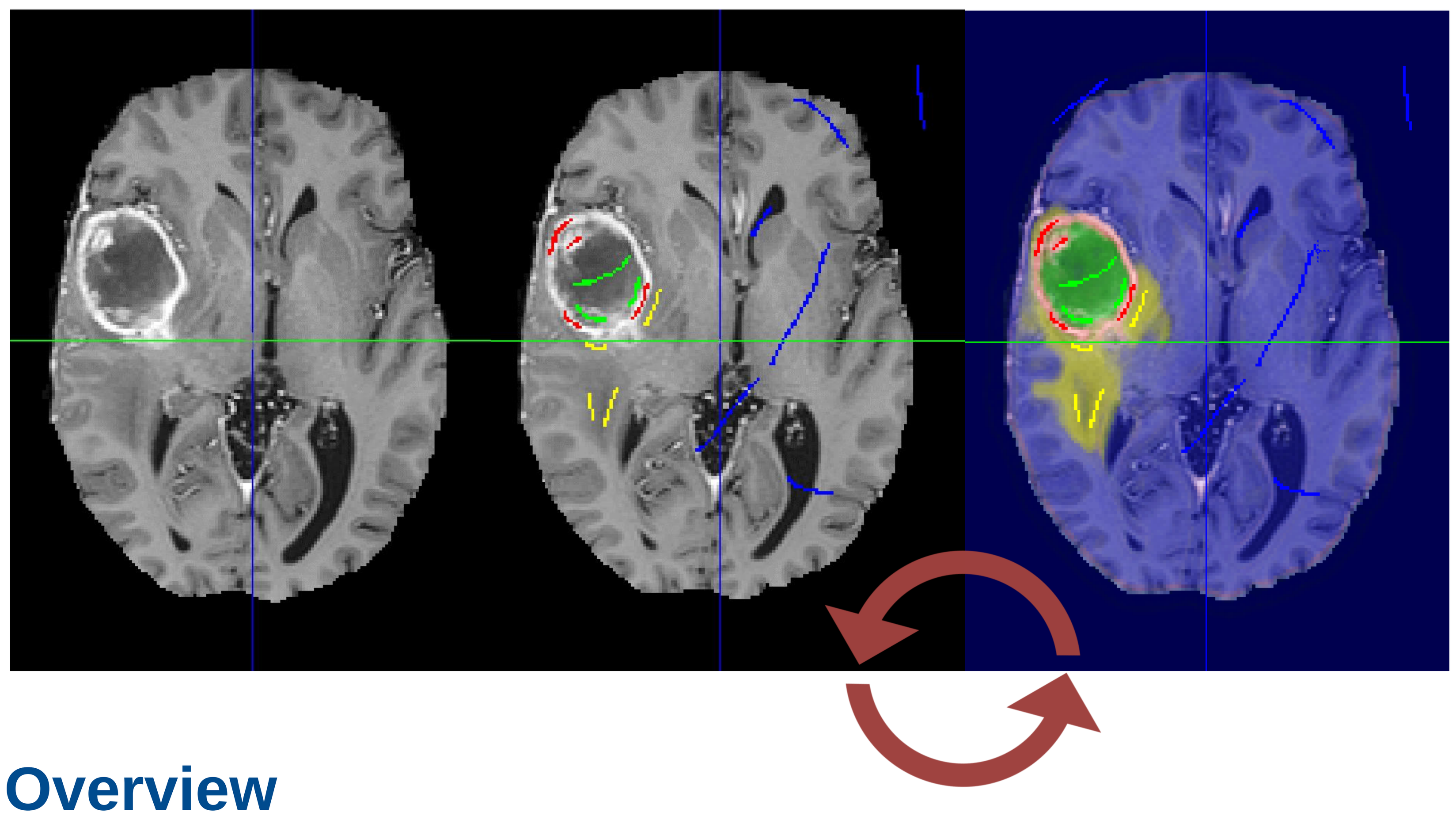
Quantification of Guidance Strategies in Online Interactive Semantic Segmentation of Glioblastoma MRI

Jens Petersen[1,2], Sabine Heiland[1], Martin Bendszus[1], Jürgen Debus[3,4], Klaus H. Maier-Hein[2]
[1]Dept. of Neuroradiology, Heidelberg University Hospital, Heidelberg, Germany
[2]Junior Group Medical Image Computing, German Cancer Research Center (DKFZ), Heidelberg, Germany
[3]Dept. of Radiation Oncology, Heidelberg University Hospital, Heidelberg, Germany
[4]Clinical Cooperation Unit Radiation Oncology, German Cancer Research Center (DKFZ), Heidelberg, Germany

dkfz. GERMAN CANCER RESEARCH CENTER IN THE HELMHOLTZ ASSOCIATION
Research for a Life without Cancer

Interactive segmentation promises to combine the speed of automatic approaches with the reliability of manual techniques. Its performance, however, depends largely on live iterative inputs by a human supervisor. For the task of glioblastoma segmentation in MRI data using a Random Forest pixel classifier we quantify the benefit in terms of speed and segmentation quality of user inputs in falsely classified regions as opposed to guided annotations in regions of high classifier uncertainty. The former results in a significantly higher area under the curve of the Dice score over time in all tumor categories. Exponential fits reveal a significantly higher final Dice score for larger tumor regions (gross tumor volume and edema) but not for smaller regions (necrotic core, non-enhancing abnormalities and contrast-enhancing tumor). Time constants of the exponential fits do not differ significantly.



Overview

- Interactive, iterative, scribble-based segmentation
- Compare corrective annotations and guided annotations (highlighting uncertain regions)

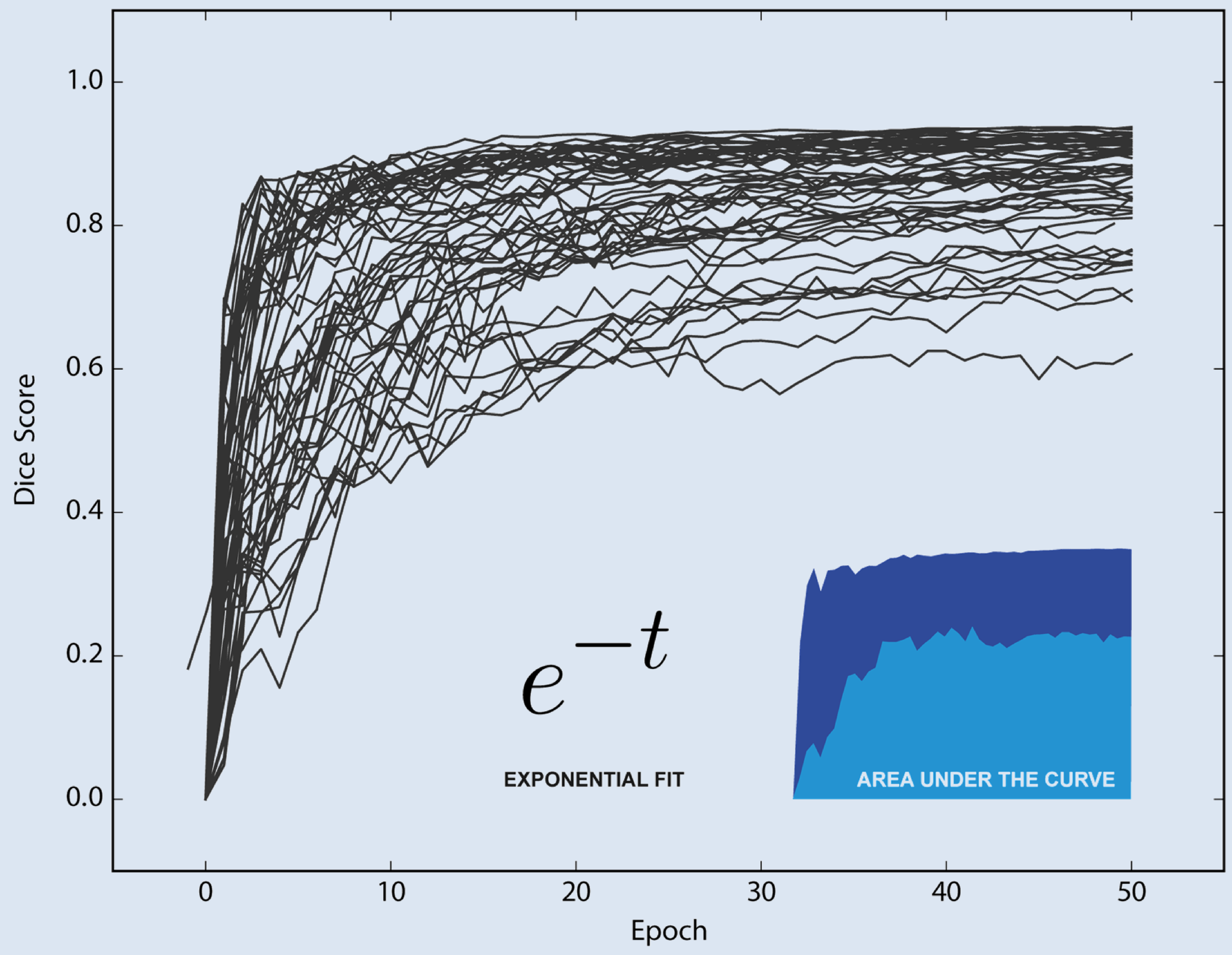
Methods & Data

- BraTS 2013 Glioblastoma MRI data
- Random Forest (50 trees, 10 maximum depth, Gini impurity splits)
- Measure Dice score over number of interactions
- Evaluate area under the curve AUC and both time constant and infinity constant of exponential fit
- Compare methods with paired T-test

Interaction Modes

$$d(t) = d_\infty \cdot \left(1 - \exp^{-t/\tau}\right)$$

- Uncertainty-guided
  Annotate in region of highest classifier uncertainty
- Random Corrections
  Annotate where classifier is wrong
- Balanced Corrections
  Annotate where classifier is wrong, balancing inputs among classes



Uncertainty-guided vs Random Corrections

		GTV	Core	Edema	Non-Enhanc.	Enhanc.
	Δ	**−0.056(56)**	0.014(118)	**−0.078(87)**	−0.014(87)	−0.004(99)
AUC	t	**−4.476**	0.508	**−3.922**	−0.720	−0.194
	p	**< 0.001**	0.617	**< 0.001**	0.480	0.848
	Δ	−0.002(39)		−0.016(41)		0.028(92)
τ	t	−0.195		−1.598		1.170
	p	0.848		0.130		0.262
	Δ	**−0.040(89)**		**−0.065(69)**		**0.015(27)**
d_∞	t	**−4.018**		**−3.805**		**2.144**
	p	**0.002**		**0.002**		**0.006**

Uncertainty-guided vs Balanced Corrections

		GTV	Core	Edema	Non-Enhanc.	Enhanc.
	Δ	**−0.045(62)**	**−0.105(67)**	−0.057(88)	**−0.094(72)**	**−0.076(107)**
AUC	t	**−3.179**	**−6.806**	−2.833	**−5.718**	**−3.130**
	p	**0.005**	**< 0.001**	0.011	**< 0.001**	**0.006**
	Δ	−0.010(37)		−0.024(35)		0.040(89)
τ	t	−1.081		−2.723		1.185
	p	0.297		0.016		0.089
	Δ	**−0.034(26)**		**−0.048(49)**		−0.001(18)
d_∞	t	**−5.095**		**−3.971**		−0.142
	p	**< 0.001**		**0.002**		0.889

Balanced Corrections vs Random Corrections

		GTV	Core	Edema	Non-Enhanc.	Enhanc.
	Δ	**0.012(15)**	**−0.118(116)**	**0.022(17)**	**−0.080(57)**	**−0.072(85)**
AUC	t	**3.509**	**−4.447**	**5.773**	**−6.150**	**−3.701**
	p	**0.003**	**< 0.001**	**< 0.001**	**< 0.001**	**0.002**
	Δ	−0.012(40)		−0.015(43)		0.032(67)
τ	t	−1.216		−1.551		1.858
	p	0.241		0.137		0.083
	Δ	0.005(19)		0.010(29)		−0.013(23)
d_∞	t	1.143		1.595		−2.346
	p	0.269		0.127		0.033

Key Findings

1. Corrective annotations better than annotations in regions of high classifier uncertainty
2. Important to balance inputs when class imbalance is large
3. Significant differences only in segmentation quality, not speed

Related: Petersen et al. „Effective User Guidance in Online Interactive Semantic Segmentation", Proc. SPIE Medical Imaging, 2017