Supplementary materials for "Menze et al. *A generative probabilistic model and discriminative extensions for brain lesion segmentation - with application to tumor and stroke.* IEEE Transactions on Medical Imaging 2016."

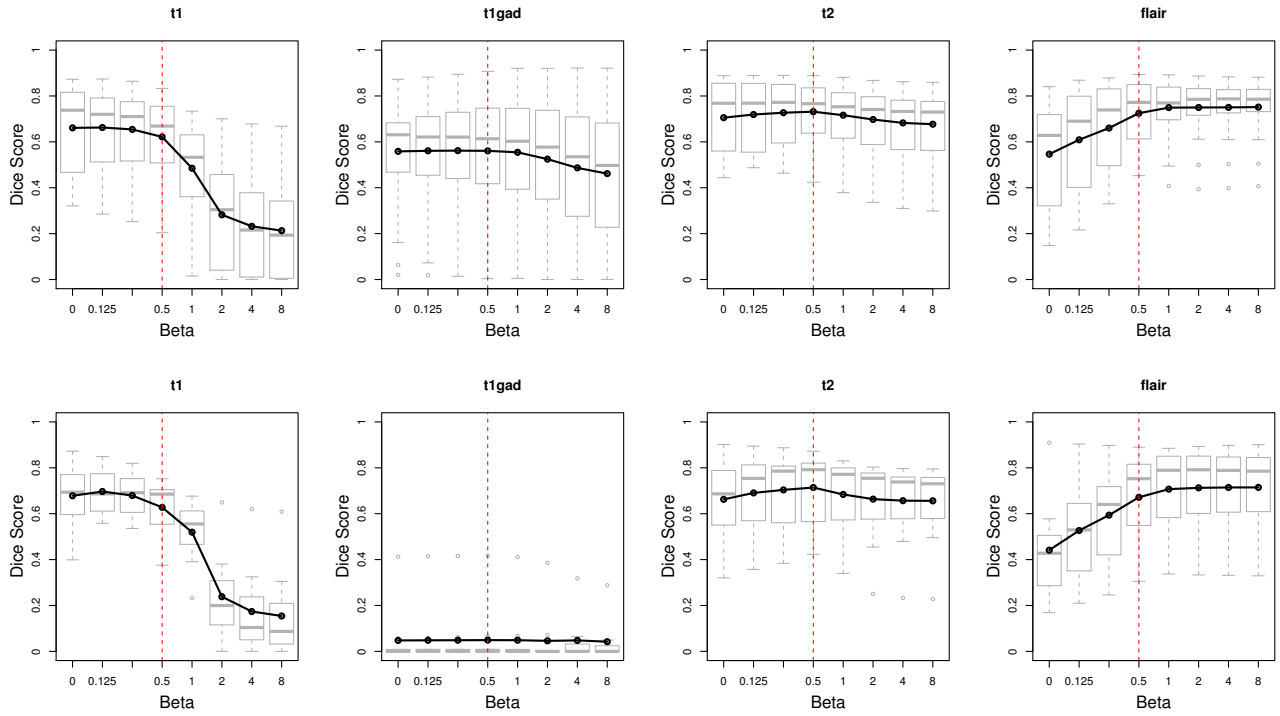

Fig. 7. Optimal spatial regularization of the tumor prior: high-grade (top row), and low-grade cases (bottom row). Reported are Dice scores for channel-specific segmentations for low- and high-grade cases of the BRATS training set, testing different values of regularization parameters $\beta \in [0, .125, ..., 8]$ in Eq. (9). Gray boxplots represent quartiles, with notches indicating outliers. Black lines and circles correspond to the mean performance. While T1c and T2 segmentations are rather insensitive to the choice of β, we choose an intermediate value of $\beta = .5$ (red dashed line) in further experiments.

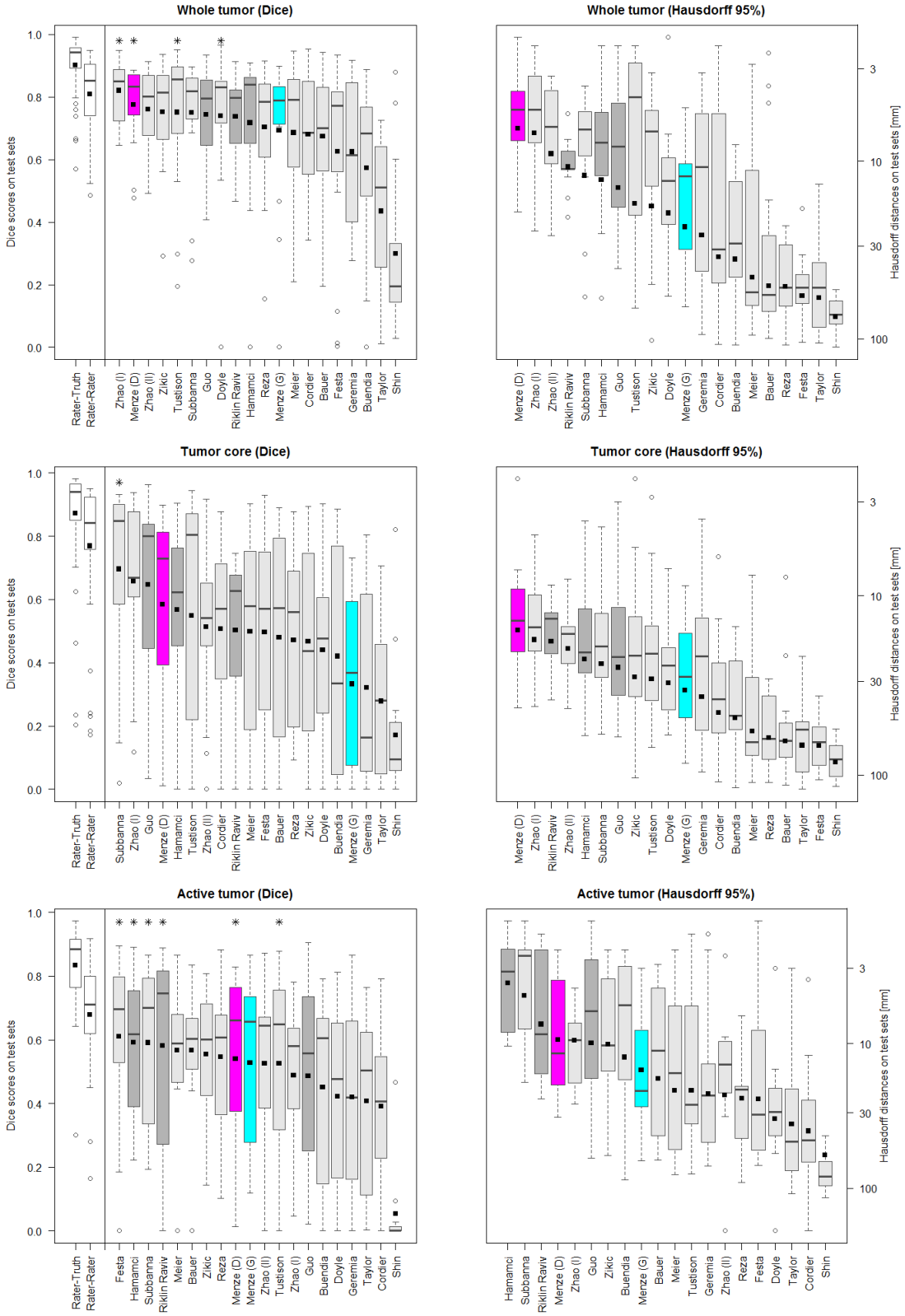

Fig. 8. Results from the BRATS evaluation reporting Dice scores and Hausdorff distances for the 'off-site' test (see Figure 8 in [1]). Methods are ranked according to the average Dice score and the robust Hausdorff distance and boxplots indicate quartiles and outliers. Results of the generative model with the two discriminative model extensions are shown in magenta and cyan, the first corresponding to results of the voxel-wise discriminative-model (*Menze (D)*) and those that have been generated by removing false positive regions from the segmentations of the generative model (*Menze (G)*). Competing inter-active segmentation methods are indicated by dark gray boxes. White boxplots report the inter-rater Dice scores, with scores calculated between individual raters (*Rater-Rater*) and between the consensus and raters (*Rater-Truth*). Stars on top of the boxes indicate methods with results that do not differ significantly from the inter-rater variation ($p<.05$). Also refer to the BRATS evaluation paper [1] for additional details (http://dx.doi.org/10.1109/TMI.2014.2377694).