

## Evaluation of Submitted Results

Vladimir Pekar

June 11, 2009

The evaluation of the submitted results will be done according to the following criteria:

### 1. Hausdorff distance.

The Hausdorff metric measures the maximum distance of a point in a set to the nearest point in the other set:

$$d_H(X, Y) = \max\{\max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y)\}.$$

This distance will only be evaluated in axial slices where the expert manual delineations are present. Its large value gives an indication that automated segmentation is not accurate and manual corrections are required in this particular slice. Since deviations less than 3mm are often considered as acceptable by the clinicians, an additional evaluation criterion will be the number of slices per dataset with the Hausdorff distance exceeding 3mm. The average distance over all slices will also be used to produce the final score.

The technical implementation of the Hausdorff metric for the challenge is done by computing a Euclidean distance map around the binary mask, where its interior is considered to have the distance value of 0. The resulting Hausdorff distance value of -1 for a particular axial slice means that no automatically generated delineation is present in a slice where a manual expert delineation exists.

### 2. Dice similarity coefficient.

This criterion is valued from 0 and 1, and is a measure of volumetric overlap computed as:

$$\kappa = 2 \times \frac{|X \cap Y|}{|X| + |Y|},$$

where  $|\cdot|$  is the number of pixels/voxels. Analogously to the Hausdorff distance, this criterion will also be only evaluated in axial slices where the manual delineations are present, and the average will be used to produce the final score.

A Windows executable program is available to the participants for download to test their results against the expert manual delineations provided in the training data. The program syntax is:

```
maskcompare filename_reference filename_mask > statistics.dat
```

Please allow for approx. 5sec processing time before opening the output file. The produced statistics in the output file can be easily analyzed and plotted by using some standard data analysis software.