Introducing Dice, Jaccard, and Other Label Overlap Measures To ITK

Nicholas J. Tustison, and James C. Gee

December 1, 2009

Penn Image Computing and Science Laboratory
University of Pennsylvania

Abstract

Although the KappaStatisticImageToImageMetric can be used to obtain the Dice metric (or mean overlap), there are other related measurements that are useful for evaluating results derived from various image analysis tasks. These measures include the target overlap, union overlap (or Jaccard coefficient), and false positive/negative errors. There are also related measures for multilabeled images. The class included with this submission, the `itk::LabelOverlapMeasuresImageFilter`, is meant to provide an easy mechanism for calculating such measures.

1 Introduction

Many of the details discussed here are explained in a recent evaluation study of 14 image registration algorithms by Klein et al. [1] and the references contained therein. Please see those references for further information.

Given a source image and a target image, these set of overlap measurements include the following (please see Figure 1 for a pictorial representation of the various measure components):

- Overlap Agreement Measures
  - Target Overlap (for individual labeled region, r)

$$TO_r = \frac{|S_r \cap T_r|}{|T_r|} \tag{1}$$

  - Total Overlap (for all regions)

$$TO = \frac{\sum_r |S_r \cap T_r|}{\sum_r |T_r|} \tag{2}$$

  - Union Overlap or Jaccard coefficient (for individual labeled region, r)

$$UO_r = 2\frac{|S_r \cap T_r|}{|S_r \cup T_r|} \tag{3}$$

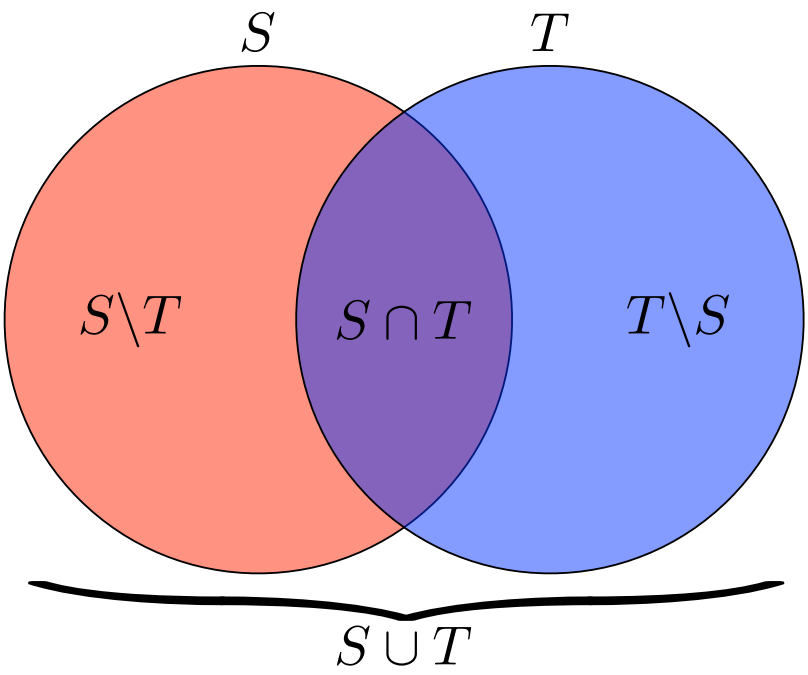

Figure 1: Schematic Venn diagram for understanding the various measure components used to calculate the overlap agreement measures and the overlap error measures.

- Union Overlap or Jaccard coefficient (for all regions)

$$UO = 2\frac{\sum_r |S_r \cap T_r|}{\sum_r |S_r \cup T_r|} \tag{4}$$

- Mean Overlap or Dice coefficient (for individual labeled region, r)

$$MO_r = 2\frac{|S_r \cap T_r|}{|S_r| + |T_r|} \tag{5}$$

$$= \frac{2 \times UO_r}{1 + UO_r} \tag{6}$$

- Mean Overlap or Dice coefficient (for all regions)

$$MO = 2\frac{\sum_r |S_r \cap T_r|}{\sum_r |S_r| + |T_r|} \tag{7}$$

$$= \frac{2 \times UO}{1 + UO} \tag{8}$$

- Overlap Error Measures
  - False Negative Error (for individual labeled region, r)

$$FN_r = \frac{|T_r \backslash S_r|}{|T_r|} \tag{9}$$

(a) (b)

Figure 2: (a) Source and (b) target image for the 2-D example accompanying this submission.

– False Negative Error (for all regions)

$$FN = \frac{\sum_r |T_r \setminus S_r|}{\sum_r |T_r|} \tag{10}$$

– False Positive Error (for individual labeled region, r)

$$FP_r = \frac{|S_r \setminus T_r|}{|S_r|} \tag{11}$$

– False Positive Error (for all regions)

$$FP = \frac{\sum_r |S_r \setminus T_r|}{\sum_r |S_r|} \tag{12}$$

Being inspired by the `itk::LabelStatisticsImageFilter`, the proposed filter behaves similarly (e.g. is multithreaded).

2 Example

The test included with this submission prints an organized dump to the screen of all the measures discussed previously. We include two images `sourceImage.nii.gz` and `targetImage.nii.gz` which produces the following result.

```
[15:31:46] $ ./itkLabelOverlapMeasuresImageFilterTest 2 sourceImage.nii.gz targetImage.nii.gz
                            ************  All Labels *************
                      Total  Union (jaccard)      Mean (dice)   False negative   False positive
                   0.250444         0.113433         0.203753         0.749556         0.828264
                       ************ Individual Labels *************
Label                Target  Union (jaccard)      Mean (dice)   False negative   False positive
    1              0.706131         0.319312         0.484058         0.293869         0.631753
    2              0.158291        0.0609874         0.114964         0.841709         0.909742
    3                     0                0                0                1                1
    4                   nan                0                0              nan                1
    5              0.258157         0.155852         0.269674         0.741843         0.717733
    6              0.247312         0.174715          0.29746         0.752688         0.626883
```

References

[1] Arno Klein, Jesper Andersson, Babak A Ardekani, John Ashburner, Brian Avants, Ming-Chang Chiang, Gary E Christensen, Louis D Collins, James Gee, Pierre Hellier, Joo Hyun Song, Mark Jenkinson, Claude Lepage, Daniel Rueckert, Paul Thompson, Tom Vercauteren, Roger P Woods, J. John Mann, and Ramin V Parsey. Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *Neuroimage*, Jan 2009. 1